

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 16, 2004, 17:52:09 ; Search time 7695 Seconds  
(without alignments)  
4156.876 Million cell updates/sec

Title: US-10-049-957-4  
Perfect score: 3936  
Sequence: 1 MRGPGALWLLALRTVLGG.....APLLPLLLPALAARLLPPAL 738

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10049957/runat\_14052004\_100249\_1257/app\_query\_fasta\_1.903  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=500 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10049957@cgn2\_1.3284 @runat\_14052004\_100249\_1257 -NCPU=6 -ICPU=3  
-NO MWAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_on.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3936	100.0	2368	6	A00127 H. sapiens m
2	3936	100.0	2368	6	AR084661 Sequence
3	3936	100.0	2368	6	E01410 cDNA encodi
4	3936	100.0	2368	6	AR232475 Sequence
5	3936	100.0	2368	6	AR364343 Sequence
6	3936	100.0	2368	6	BD093599 Chondroge
7	3936	100.0	2368	6	BD107341 Novel cho
8	3936	100.0	2368	6	BD107341 Novel cho
9	3834	97.4	2154	6	BD012530 Method fo
10	3441	87.4	2388	6	BD093598 Chondroge
11	3441	87.4	2388	6	BD107340 Novel cho
12	3437	87.3	2388	4	AB010995 Oryctolag
13	3300	83.8	4068	6	AX212239 Sequence
14	3300	83.8	4102	10	BC040347 Mus muscu
15	3300	83.8	4158	6	BD093609 Chondroge
16	3300	83.8	4158	10	AB024336 Mus muscu
17	2493.5	63.4	3547	5	GGEOS47GN
18	1288.5	32.7	2444	5	AY335444 Pagrus ma
19	1274.5	32.4	2411	4	CHLACF
20	1272.5	32.3	2557	5	SMOTRFR
21	1271	32.3	2437	5	D89088
22	1270.5	32.3	2298	9	HUMTF
23	1270.5	32.3	2326	4	CHU53857
24	1270.5	32.3	2327	6	AR087435 Sequence
25	1270.5	32.3	2347	6	AX410394 Sequence
26	1270.5	32.3	2347	9	S95936 transferrin
27	1270	32.3	2423	9	HSU07643 Human lacto
28	1269.5	32.3	2097	6	BD105975 Animal mo
29	1269.5	32.3	4896	9	HUMLACFE
30	1268.5	32.2	2330	9	BC059367 Homo sapi
31	1268	32.2	2311	9	HSLTR
32	1267	32.2	2297	9	HUMLF2A
33	1265.5	32.2	2357	4	BOVLACFERR
34	1264.5	32.1	2379	9	BC022347 Homo sapi
35	1264	32.1	2437	5	D89091
36	1263.5	32.1	2231	4	ECA010930 Equus Cab
37	1263.5	32.1	2302	4	ECU21127 Equus Cabal
38	1263.5	32.1	2310	4	HRSTFRA
39	1263	32.1	2360	6	AR066024 Sequence
40	1263	32.1	2360	6	AR149765 Sequence
41	1263	32.1	2360	6	I28419 Sequence 1
42	1263	32.1	2360	6	I28532 Sequence 1
43	1263	32.1	2360	6	AR410144 Sequence
44	1263	32.1	2360	9	HUMLACTOPE
45	1263	32.1	2431	5	D89089 Salvelinus

ALIGNMENTS

RESULT 1

A00127  
LOCUS A00127 2368 bp RNA linear PAT 28-JAN-1993  
DEFINITION H.sapiens mRNA for melanoma associated antigen P97.  
ACCESSION A00127  
KEYWORDS melanoma associated protein; p97 melanoma-associated antigen.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2368)  
AUTHORS Brown, J.P., Plozman, G.D., Hellstrom, K.E., Purchio, A.F.,  
Pennathur, S., Estlin, C.D., Rose, T.M., Hellstrom, I. and Hu, S.L.  
JOURNAL Patent: GB 2188637-A 1 07-OCT-1987;  
Oncogen

FEATURES  
source  
1..2368  
/organism="Homo sapiens"  
/mol\_type="unassigned RNA"  
/db\_xref="taxon:9606"  
61..2277  
/codon\_start=1  
/product="melanoma associated antigen P97"  
/protein\_id="CAA00012.1"  
/db\_xref="GI:14413"  
/db\_xref="GOA:P08582"  
/db\_xref="SWISS-PROT:P08582"  
/translation="MRGPGALWLLLRITVLGVMEVRWCATSDPEQHKCNMGSEARF  
EAGIQPSLLCVRRGSDHCQVQIAAQADAITLDGGAIIYEAGKEHGLKPVVGEYDDE  
VGYTYAVAVVRSSHTVITDLKGVKSGHTINRTGVNVPVGLVSRGLSVNGCDV  
LKAVSDYFPGSCVPGAGETSYESICRLCRGDSGEGVCDKSPLERYDYSGAFRCLA  
EGAGDVFVKGSTVLENTDGTLLPSWGOALLSQDFELLCRDGSRADYVTEWROCHLAV  
PAHAVVRADTDGGLIIRLNEGORLSHEGSSFFQMPSSSEAYGQKDLIFKDSISELVP  
IATQYEWLGEHMRKMGQLCDPNRPYLRWCVLSTPEIQGDMVAFRQRLK  
PEIQCVSAKSPHCHERIQAEQVDVLTLSGEDIYTAGKKYGLVPAAGEHVPEDSSNS  
YVYVAVRRDSSHVAITDELGRKRSCHAGFGSPAGWDVPVGGALIRQFIRPKDCDLVT  
AVSFFFNASCVSNPNKYPSLICALCVDEQGRNKCVCNISOERYGVGRGAFRCLVEN  
AGDAVFVHTTVFNTGNHSEPWAAELRSEDYELLCPNGARAEVSPFAACNLAIQIP  
HADVVRPDNTFTVYGLLDKAQDLFGDDHKNKGKMFEDSSNYHQQDLIFKDATVRAP  
VGEKTYRGLWGLDYVALEGNSSQCSGAAAPAPGAPLLPLLLPALAARLLPPAL"

ORIGIN  
Alignment Scores:  
Pred. No.: 3,488-286 Length: 2368  
Score: 3936.00 Matches: 738  
Percent Similarity: 100.00% Conservatve: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-049-957-4 (1-738) x A00127 (1-2368)

Qy 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20  
Db 61 ATGGGGGTCGAGCGGGCTCTGTGGCTCTCTCTGGCTCTGGCCACCGTCTCGGAGGC 120  
Qy 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
Db 121 ATGGAGGTGCGGTGGTGGCCACCTCGGACCCAGAGCAGCAAGTCGCGCAATGAGC 180  
Qy 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
Db 181 GAGGCTTCCGGGAAGCGGGATCCAGCCCTCCCTCTCTGGTCCGGGGCACCTCCGCC 240  
Qy 61 AspHisCysValGlnLeuIleAlaGlnGlnAlaAspAlaIleThrLeuAspGlyGly 80  
Db 241 GACCACCTCGCTCCAGCTCATCGCGCCCGAGGAGGCTGACGCGCATCACTCTCGATGGAGGA 300  
Qy 81 AlaIleThrGluAlaGlyLysGluHisGlyLysProValValGlyGluValThrAsp 100  
Db 301 GCCATCTATGAGCGGGGAAGAGGACGCGCTGAAGCGGGTGGTGGGCAAGTGTACGAT 360  
Qy 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValArgSerSerHisValThr 120

361 CAAGAGTCGGTACTCTTATACGCGTGGCTGTGTGTCAGGAGAGCTCCCATGTGACC 420  
121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp 140  
421 ATTACACCCCTGAAGCGGTGAAGTCTCTGCCACACCGGGCATCAATCGCACAGTGGGTGG 480  
141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
481 AACGTGCGCGTGGCTACCTGGTGAGAGCGCGCCCTCTCGGTGATGGGTGGTGCATGTA 540  
161 LeuLysAlaValSerAspTyrPheGlyLysCysValProGlyAlaGlyGlyThrSer 180  
541 CTCAAGAGCTGACGAGCATTTTGGGGGCGAGCTGCTCCCGGGGCGAGAGACCACT 600  
181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200  
601 TACTCTGAGTCCCTCTCTCGCTCTGAGGGGTGACAGCTCTTGGGGGAGGGGTGTGTAC 660  
201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGlyGly 220  
661 AAGAGCCCTCTGGAGAGATCTACGCTACAGCGGGGCTTCCGCTGCTGGCGAAGGG 720  
221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
721 GCAGGGGACGTGGCTTTGTGAGCACACGACGCTACTGGAGAACACGGATGGGAAGACG 780  
241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260  
781 CTTCCTCTCTGGGGCCAGGCCCTCTCTGTCACAGACTTCGAGCTGCTGTGCGGGATGGT 840  
261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
841 AGCGGGCCGATGTCACGAGTGGAGCAGTGCCTCTGCGCCGGTGCCTGCTTCACGCC 900  
281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300  
901 GTGGTGTGCGGGCCGACACAGATGGGGCTCATCTTCGGCTGCTCAACGAAGGCCAG 960  
301 ArgLeuPheSerHisGlyLysSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
961 CGTCTGTTCAGCCACGAGGCGAGCTTCAGATGTTGAGTCTGAGGCTATGGCCAG 1020  
321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
1021 AAGGATCTACTCTTCAAGACTCTACCTCGAGCTTGTGCCATCGCCACACACACCTAT 1080  
341 GluAlaTrpLeuGlyHisGlyTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360  
1081 GAGCGGTGGCTGGGCCATGAGTACCTGCACGCCATGAAGGTCTGCTCTGTGACCCCAAC 1140  
361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
1141 CGGTGCCCCCTACCTCGCGTGGTGTGTCTCTCCACTCCCGAGATCCAGAAATGTGGA 1200  
381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
1201 GACATGCGGTGGCTTCCCGCGGCGAGCTCAAGCCAGAGATCCAGTGGCTGTGAGCC 1260  
401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrIleu 420  
1261 AAGTCCCCCAACTGCTGAGCGGATCCAGGCTGAGCAGGTCGAGCTGTGACCTA 1320  
421 SerGlyGluAspIleThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
1321 AGTGGCAGGACATTTACCGCGGGGAAGAAGTACGCGCTGCTGCTCCCGCAGCGGGAG 1380  
441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgAsp 460  
1381 CACTATGCGCGGAAGACGACGCACTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480



QY 301 ArgLeuPheSerHisGluCysSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
DB 961 CGTCTGTTTCAGCAGGCGGAGAGCTTCAGATGTTTCAGCTTCGAGGCTATGGCCAG 1020  
QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
DB 1021 AAGGATCTACTCTTCAAGACTCTTACCTCGGAGCTTGTGCCCATCGCCACACAGACTAT 1080  
QY 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360  
DB 1081 GAGGCGTGGCTGGGCGCATGAGTACCTGCAGCGCATGAAGGGTCTCTCTGTGACCCCAAC 1140  
QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
DB 1141 CGGCTGCCGCCCTACTCGGCTGTGTGTGTCTCTCCACTCCCGAGATCGAAGTGTGA 1200  
QY 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
DB 1201 GACATGGCGGTGGCTTCGCGCGGAGCGCTCAAGCCAGAGATCCAGTGGTGTGACGC 1260  
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
DB 1261 AAGTCCCCCACACTGCTGATGAGCGGATCCAGGCTGAGCAGGTGCGCTGTGACCCCTA 1320  
QY 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
DB 1321 AGTGGCGAGGACATTTACGCGCGGGAAGATGAGGCTGTGTTCGCGAGCGCGGCGAG 1380  
QY 441 HistYrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValAlaValArgArgAsp 460  
DB 1381 CACTATGCCCGGAAGACAGCAGCAACTCGTACTAGTGGTGGCGGTGCTGCCAGCGGAC 1440  
QY 461 SerSerHisAlaPheThrLeuAspGlnLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
DB 1441 AGCTCCACGCGCTTCACCTTGTGATGAGCTTCGGGCGAAGCGCTCTGCCAGCGCGGTTTC 1500  
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
DB 1501 GGCACCTTCGAGGTGGGATGTCCTCGGTGGTGCCTTATTCAGAGAGGCTTCACTCGG 1560  
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
DB 1561 CCCAAGGACTGTGACGTCTCACAGCAGTGAGCGAGTCTTCAATGTCAGCTGCGTGCC 1620  
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGlnGln 540  
DB 1621 GTGAACAAACCCCAAGAACTACCCCTCTCGCTGTGTGCACTGTGCGTGGGCGACGAG 1680  
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
DB 1681 GGCCGCAACAGTGTGTGGGCAACGCGGAGCGGATACGGCTACCGCGGCGGCTTC 1740  
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
DB 1741 AGGTGCTGTGTGAGAATCGGGTGACGTGCTTGTGTCGTCAGGCACACACCGTCTTTGAC 1800  
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600  
DB 1801 AACAAACAGCGGCACAAATTCGAGCGCTGGGCTGTGAGCTCAGGTGAGGACTATGAA 1860  
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
DB 1861 CTGCTGTGCCCCAACGCGGCGCCGAGCGGAGGTGCCATTTGAGCTGCGAACCTGGCA 1920  
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
DB 1921 CAGATACCAACCCACGCGCTGATGTTCGCGCCCGACACCAACATCTTCCCGTGTATGA 1980  
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660  
DB 1981 CTGTGGGACAGGCGCCAGGACCTGTGTGGAGACGACCAATTAAGACGGGTCAAAATG 2040  
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla 680

DB 2041 TTGACTCTCTCAACTATCATGGCCAGACCTGCTTTCAAGGATGCCACGTCGCGGCG 2100  
QY 681 ValProValGlyGluLysThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaAla 700  
DB 2101 GTGCTGTGCGAGAGAAAACCACTACCGCGCTGGCTGGGCTGGACTACGTGGCGCG 2160  
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720  
DB 2161 CTGGAAGGATGTGCTCTCAGAGTGTCTGGCGGCGAGCGGCCCGCGCGCGCGGCC 2220  
QY 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738  
DB 2221 CTGCTCCGCTGCTGCTGCCCGCCCTCGCGCGCGCTGCTCCGCGCGCGCTC 2274  
RESULT 3  
E01410  
LOCUS  
DEFINITION cdna encoding polypeptide having the melanoma-associated p97  
antigen.  
ACCESSION E01410  
VERSION E01410.1 GI:2169666  
KEYWORDS JP 1987294698-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 2368)  
AUTHORS JiYosefu,P.B., ChiYaaruZu,D.I., Guregorii,D.P., Teimoshii,E.R.,  
karuru,I.H., Ingegarudo,H., Ansonii,E.P., Shiu,R.F. and  
Suridoharu,P.  
TITLE VACCINE AGAINST MELANOMA  
JOURNAL Patent: JP 1987294698-A 1 22-DEC-1987;  
COMMENT ONKOOGEN  
OS homo sapiens  
PN JP 1987294698-A/1  
PD 22-DEC-1987  
PF 06-FEB-1987 JP 1987026191  
PR 07-FEB-1986 US 86 827313, 26-JAN-1987 US 87 7230 PI  
JIYOSFUFU PII BURAUIN, CHIYAARUZU DEI IISUTEIN, PI GUREGORII DEI  
PUREOMAN,  
PI TEIMOSHII EMU ROOZU, KARURU II HERUSUTOROMU,  
PI INGEARUDO HERUSUTOROMU,  
PI ANSONII EFU PRACHIO, SHIU ROKU FUJ, SURIDOHARU PENNATEYUURU PC  
C07K13/00,A61K39/00,A61K39/12,C12N1/20,C12N5/00,C12N7/04, PC  
C12N15/00,  
PC C12P21/02, (C12N1/20,C12R1:19), (C12N5/00,C12R1:91), (C12P21/02,  
C12R1:01),  
PC (C12P21/02,C12R1:645), (C12P21/02,C12R1:91);  
CC strandedness: Single;  
CC topology: linear;  
CC hypothetical: No;  
CC anti-sense: No;  
CC \*source: tissue types=melanoma;  
CC \*source: cell\_line=SK-MEL28;  
FH Key Location/Qualifiers  
FH CDS 61..2277  
FT /product='polypeptide having the FT  
FT melanoma-associated p97  
FT antigen'  
FT mat\_peptide 61..2274  
FT /product='polypeptide having the FT  
FT melanoma-associated p97  
FT antigen'  
FT 5'UTR <1..60  
FT 3'UTR 2278..>2308.  
FEATURES  
source  
1..2368  
/organism="Homo sapiens"  
/mol\_type="genomic RNA"  
/db\_xref="taxon:9606"  
ORIGIN



## Alignment Scores:

Pred. No.: 3,48e-286 Length: 2368  
Score: 3936.00 Matches: 738  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0

US-10-049-957-4 (1-738) x E01410 (1-2368)

```
QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuLeuAlaLeuArgThrValLeuGlyGly 20
DB 61 ATGCGGGGTCGAGCGGGGCTCTGTGGCTGCTCTGGCTCTGGCAACCGTGTGCGAGGC 120
QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40
DB 121 ATGGAGTGGGTGGTGGTGGCCACCTCGGACCCAGACGACACAGTGGCGGCAATGAGC 180
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60
DB 181 GAGGCCCTTCGGGAAGCGGGCATCCAGCCCTCCCTCTCGTCCGGGGCACTCCGCC 240
QY 61 AspHisCysValGlnLeuIleAlaGlnGlnAlaAspAlaIleThrLeuAspGlyGly 80
DB 241 GACCACTGGCTCAGCTATCGGGCCAGGGGCTGAGCGCCATCATCTTGGATGGAGGA 300
QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100
DB 301 GCATCTATAGCGGGGAAGAGACACGGCCCTGAGCCGGTGGTGGCGCAAGTGTACGAT 360
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValAlaValArgSerSerHisValThr 120
DB 361 CAAGAGTCCGTACCTCTATTACGCCGTGGTGTGGTCCAGGAGAGCTCCCATGTGACC 420
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp 140
DB 421 ATTGACACCTCTGAAGCGGTGAAGTCTTCCACACGGGCATCAATCGCACAGTGGGCTGG 480
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160
DB 481 ACAGTCCCGTGGGTCTACTGGTGGAGACGGCCGCTCTCGGTGATGGGCTGCCATGTA 540
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180
DB 541 CTCAAGCTGTGAGCACTATTTTGGGGCAGCTGGTCCCGGGGCGAGAGACCACT 600
QY 181 TyrSerGluSerLeuLysCysArgLeuLysCysArgGlyAspSerSerGlyGlyValCysAsp 200
DB 601 TACTCTGAGTCCCTCTGTGCGCTCTGACGGGTGACAGCTCTGGGGAAGGGTGTGTGAC 660
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyValaPheArgCysLeuAlaGlyGly 220
DB 661 AAGAGCCCTCTGGAGAGATCTACGACTACGCGGGGCTTCGGTGGCTTGGCGGAGGG 720
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240
DB 721 CGAGGGGACGTGGCTTTTGTGAAGCACACACCGGTACTGGAAACACCGATGGGAAGCG 780
QY 241 LeuProSerTrpGlyClnAlaLeuLeuSerGlnAspPheGluLeuLeuLysCysArgAspGly 260
DB 781 CTTCCCTCTGGGGCCAGGCCCTGTGTGCACAGGACTTCAGCTGTGTGGCGGATGGT 840
QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280
DB 841 AGCGGGCCGATGTACCGAGTGGAGGAGTGGCCATCTGGCCCGGGTGTCTGTCTACGCC 900
QY 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300
DB 901 GTGGTGGTCCGGGCCGACACAGATGGGGCCCTCATCTTCCGGCTGTCTCAACGAAGCCAG 960
QY 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320
```

```
DB 961 CGTCTGTTAGCCACGAGGGCAGCAGCTTCCAGATGTTCCAGCTCTGAGCCCTATGCCCAG 1020
QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340
DB 1021 AAGGATCTACTCTTCAAGAGACTCTTACTCGGAGCTTGTGCCATCGCCACACAGACCTAT 1080
QY 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuLysCysAspProAsn 360
DB 1081 GAGGCGTGGTGGGCCATGAGTACCTGCACGCGCATGAAGGGTCTGCTCTGTGACCCCAAC 1140
QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380
DB 1141 CGGCTGCCCCCTACTCGCTGGGTGTGTCTCCACTCCCGAGATCCAGAGGTGGA 1200
QY 381 AspMetAlaValAlaPheArgGlnArgLeuLysProGluIleGlnCysValSerAla 400
DB 1201 GACATGGCGGTGGCTTCCGCGGCGAGCGCTCAAGCCAGAGATCCAGTGGTGTGAGCC 1260
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420
DB 1261 AAGTCCCGCCAACTGCTGATGAGCGGATCCAGGCTGAGCAGGTGAGCGCTGTGACCTTA 1320
QY 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440
DB 1321 AGTGGCGAGGACATTTACCGCGGGGAAGATACGGCTGGTTCGCCGACCGCGGAG 1380
QY 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp 460
DB 1381 CACTATGCCCCGGAAGACAGCAGCAACTCGTACTAGTGGTGGTGGTGGTGGAGCGGAC 1440
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480
DB 1441 AGCTCCCAACGCTTCCCTTGGATGAGCTTCGGGGCAAGCGCTCTCTCCACCGCGGTTTC 1500
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500
DB 1501 GGCAGCCCTGCGAGGTGGGATGTCCTGGTGGGTGGCTTATTCAGAGAGCTTCATCCGG 1560
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520
DB 1561 CCCAAGGACTGTGACGTCTCACAGCAGTGAGCGAGTCTTCAATGCCAGCTCGGTGCC 1620
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuLysCysAlaLeuLysValGlyAspGluGln 540
DB 1621 GTGAACAACCCCAAGAACTACCCCTCTCGTGTGTGCACTGTGGTGGGGGACGAGCAG 1680
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560
DB 1681 GGGCCCAACAGTGTGGGGCAACGAGCAGGAGCGGTATTACGGGTACCGCGGCGCTTC 1740
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580
DB 1741 AGGTCCCTGGTGGAGAAATCGGGTGACGTTCCTTCGTTCAGGCACACACACCTCTTGAC 1800
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuLysArgSerGluAspTyrGlu 600
DB 1801 AACCAAAACCGGCCCAATTCGAGCCCTGGGCTGTGTGAGCTCAGGTTCAGAGGACTATGAA 1860
QY 601 LeuLeuLysProAsnGlyValaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
DB 1861 CTGCTGTGCCCCCAACGGGGCCGAGCGAGGTGTCCCAGTTTGCAGCTTGCAGCTTGGCA 1920
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
DB 1921 CAGATACACCCCGCGCTGATGTCGGGCCCGACACCAACATCTTCCACCGTGTATGA 1980
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660
DB 1981 CTGCTGGACAAGCGCGCAGGACCTGTTTGGAGACGACCAACAATAAGAACCGGGTTCAAA 2040
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680
DB 2041 TTCGACTCTCCAACTATCATGGCCAAAGACCTGCTTTTCAAGGATGCCACCGCTCCGGCG 2100
```

QY 681 ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700  
 Db 2101 GTGCTGTGCGAGAGAAACCACTACCGCGGCTGCTGGGCTGACTACGTGGCGGG 2160  
 QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGlyAlaPro 720  
 Db 2161 CTGGAAGGGATGTCGTCTCAGCAGTCTCGGCGCAGCGGCCCGCGCGCGCGGCC 2220  
 QY 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738  
 Db 2221 CTGCTCCCGCTGCTGTCGCCGCTCGCGCGCGCGCTGCTCCCGCGCGCCCTC 2274

## RESULT 4

LOCUS AR232475 2368 bp mRNA linear PAT 20-DEC-2002  
 DEFINITION Sequence 1 from patent US 6455494.  
 ACCESSION AR232475  
 VERSION AR232475.1 GI:27274554

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2368)

AUTHORS Jefferies,W.A., McGeer,P.L., Rothenberger,S., Food,M.R., Yamada,T.

and Kennard,M.

TITLE Use of p97 and iron binding proteins as diagnostic and therapeutic

agents

JOURNAL Patent: US 6455494-A 1 24-SEP-2002;

FEATURES

source

1..2368

/organism="unknown"

/mol\_type="mRNA"

## ORIGIN

## Alignment Scores:

Pred. No.: 3,48e-286 Length: 2368

Score: 3936.00 Matches: 738

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-049-957-4 (1-738) x AR232475 (1-2368)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuLeuAlaLeuArgThrValLeuGlyGly 20

Db 61 ATGCGGGGTCCAGCGGGGCTCTGTGGTCTCTGGCTCTCGGCACCGTGTCTCGAGGC 120

QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40

Db 121 ATGGAGGTGCGGTGTGCGCCACCTCGGACCCAGAGCAGCAAGTGGCGCAACATGAGC 180

QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60

Db 181 GAGGCTTCCGGGAAGCGGCATCAGCCCTCCCTCTCTGCTCGGGGCACCTCCGCC 240

QY 61 AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80

Db 241 GACCACTGCTCAGCTCATCGCGGCCAGGAGGTGACGCCATCACTCTCGATGGAGGA 300

QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100

Db 301 GCCATCTATGAGCGGGGAAGAGACACGCCCTGAAGCGCGGTGGTGGCGGAAGTGTACCAT 360

QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValAlaValArgSerSerHisValThr 120

Db 361 CAAGAGGTGCGTACCTCTATACCGCGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 420

QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp 140

Db 421 ATTGACACCTGAAGGGGTGAAGTCTCCACACGGGCATCAATCGCACAGTGGGCTGG 480

QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
 Db 481 AACGTGCCGTGGGTACTCTGTGGAGAGCGGCCCTCTCGTGTATGGGTGGATGTA 540  
 QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyIleThrSer 180  
 Db 541 CTCAAAGCTGTACGACGACTATTTTGGGGGAGCTGCGTCCCGGGGAGGAGAGACAGT 600  
 QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200  
 Db 601 TACTCTAGTCCCTCTGTCCGCTCTGCAGGGGTGACAGCTCTGGGGAAGGGGTGTGTGAC 660  
 QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly 220  
 Db 661 AAGAGCCCTTGGAGAGATCTACACTACAGCGGGGCTTCCGGTGTCTGGCGGAAGGG 720  
 QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
 Db 721 GCAGGGGACGTGGCTTTTGTGAGCACACAGCAGGTACTGGAGAACACCGATGGGAAGACG 780  
 QY 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260  
 Db 781 CTTCCCTCTCGGGCCAGGCCCTGTGTACAGGACTTTCGAGCTCTGTCCGGATGGT 840  
 QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
 Db 841 AGCGGGCGGATGTACCGAGTGGAGGAGTGCATCTGGCCCGGGTGCCTGTCTCACGCC 900  
 QY 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300  
 Db 901 GTGGTGGTCCGGGCCGACACAGATGGGGGCTCTCATCTTCCGGCTGTCAACGAAGGCCAG 960  
 QY 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
 Db 961 CGTCTGTTCCAGCCACGAGGCGAGCAGCTTCCAGATGTTTCAGCTCTGAGCGCTATGCCAG 1020  
 QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
 Db 1021 AAGGATCTACTCTTCAAAGACTCTACTCTGGAGCTTGTGCCCATCGCCACACAGACCTAT 1080  
 QY 341 GluAlaTrpLeuGlyHisGlyTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360  
 Db 1081 GAGCGTGGCTGGGCCATGAGTACTCTGCACGCCATGAAGGGTCTGCTCTGTGACCCCAAC 1140  
 QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
 Db 1141 CGGCTGCCCTTACCTGCGTGTGTGTCTCTCCACTCCCGAGATCCAGAGTGTGGA 1200  
 QY 381 AspMetAlaValAlaPheArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
 Db 1201 GACATGGCCGTGGGCTTCCCGCGGAGGCTCAAGCCAGAGATCCAGTGTGTGAGCC 1260  
 QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
 Db 1261 AAGTCCCCCAACACTGTCATGAGCGGATCAGGCTGAGCAGGTGACGCTGTGACCCCTA 1320  
 QY 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
 Db 1321 AGTGGCGAGGACATTTACACGGCGGGGAAGTACGCGCTTCCCGCGAGCGCGCAG 1380  
 QY 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValAlaValArgArgAsp 460  
 Db 1381 CACTATGCCCCCGAAGACAGCAGCAACTCGTACTACGTGGTGGCGGTGGTGGAGCGGAC 1440  
 QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
 Db 1441 AGCTCCACGCTTCCACTTGGATGAGCTTCGGGGCAGCGCTCTCTGCCACCGCGTTTC 1500  
 QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
 Db 1501 GCGAGCCCTGAGGCTGGGATGTCCCGTGGTGGCTTATTCAGAGAGGCTTCATCCGG 1560  
 QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520



```
QY 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360
Db 1081 GAGGGTGGCTGGCCATGAGTACTGCAAGCCATGAAGGGTCTGCTCTGTGACCCCAAC 1140
QY 361 ArgLeuProTyrLeuArgTyrCysValLeuSerThrProGluIleGlnLysCysGly 380
Db 1141 CGGCTGCCCCCTACTCGCTGGTGTGTCTCTCCACTCCCGAGATCCAGAAGTGTGA 1200
QY 381 AspMetAlaValAlaPheArgGlnArgLeuLysProGluIleGlnCysValSerAla 400
Db 1201 GACATGGCCGTGGCTTCCGCGGAGCGCTCAAGCCAGAGATCCAGTGGTGTGAGCC 1260
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420
Db 1261 AAGTCCCCCAACACTGCTGATGGAGCGGATCCAGGCTGAGCGTGCAGCTGTGACCTA 1320
QY 421 SerGlyGluAspIleTyrThrAlaGlyLysIleTyrGlyLeuValProAlaAlaGlyGlu 440
Db 1321 AGTGGCAGGACATTTACACGCGCGGGAAGTACGGCTGGTTCCCGCAGCGCGGAG 1380
QY 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgAsp 460
Db 1381 CACTATGCCCGGAAGACAGACGAACTCGTACTACGTGGTGGCGCTGGTGGAGCGGAC 1440
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480
Db 1441 AGCTCCACGCTTCCCTTGGATGAGCTTCGGGCAAGCGCTCTGCACGCCGTTTC 1500
QY 481 GlySerProAlaGlyTyrAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500
Db 1501 GGCAGCCCTCGAGGCTGGGATGTCGCCGTGGTGGCTTATTCAGAGAGGCTTCATCCGG 1560
QY 501 ProlLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520
Db 1561 CCCAAGGACTGTGAGCTCTCACAGCAGTGGAGGATTTCTTCAATGCCAGTGGTGGCC 1620
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuLeuValGlyAspGluGln 540
Db 1621 GTGAACACCCCAAGAACTACCCCTCTCGTGTGTGACTGTGGTGGGGAGCAGGAG 1680
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560
Db 1681 GCGCCCAACAGTGTGTGGCAACAGCCAGGAGCGGTATTACGGCTACCGCGCGCCTTC 1740
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrValPheAsp 580
Db 1741 AGGTCCCTGGTGGAGAAATGCGGGTACGTTGCTTCGTGAGGCACACAAACCGTCTTTAC 1800
QY 581 AsnThrAsnGlyHisAsnSerGluProTyrAlaAlaGluLeuArgSerGluAspTyrGlu 600
Db 1801 AACACAAACGCCCAATTCGAGGCCCTGGGCTGTGAGCTCAGGTGAGAGGACTATGAA 1860
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
Db 1861 CTGCTGTGCCCAACAGGGGCCGAGCGAGGTGTCCAGTTTTCAGCGCTGCAACCTGGCA 1920
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
Db 1921 CAGATACACCCCGCGCTGATGTGCGGCCGAGACCAACATCTTCCCGTGTATGGA 1980
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660
Db 1981 CTGCTGGACAGGCCCGAGGACTGTGTGGAGACGACCAATAGAACGGGTTCAAATG 2040
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla 680
Db 2041 TTCGACTCTCCAACTATCATGGCCCAAGACCTGCTTTTCAAGGATGCCACCGTCCGGCG 2100
QY 681 ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700
Db 2101 GTGCTGTGGAGAGAAACCACTACCGCGGCTGGCTGGGGCTGGACTAGCTGGCGGCG 2160
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720
```

```
Db 2161 CTGGAAGGATGTGCTCTCACAGTGTCTGCGGCGCAGCGGCCCGCGGCGGCC 2220
QY 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProAlaLeu 738
Db 2221 CTGCTCCGCTGCTGCTGCGCGCCCTGCGCGCCGCTGCTCCGCGCGCCCTC 2274

RESULT 6
LOCUS BD093599 2368 bp DNA linear PAT 27-AUG-2002
DEFINITION Chondrogenesis promoter.
ACCESSION BD093599
VERSION BD093599.1 GI:22639187
KEYWORDS WO 0113951-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 2368)
AUTHORS Kato, Y. and Fujimoto, K.
TITLE Chondrogenesis promoter
JOURNAL Patent: WO 0113951-A 2 01-MAR-2001;
CHUGAI PHARMACEUTICAL CO LTD, YUKIO KATO, KATSUMI FUJIMOTO
COMMENT
OS Homo sapiens (human)
PN WO 0113951-A/2
PD 01-MAR-2001
PF 21-AUG-2000 WO 2000JP005590
PR 19-AUG-1999 JP 99P 232966
PI YUKIO KATO, KATSUMI FUJIMOTO
PC A61K45/00, A61K38/40, A61K48/00, A61K31/7088, A61K35/32, A61P19/02,
C07K14/47,
PC C07K14/79, C12Q1/02, G01N33/50, G01N33/15
CC
FH
Key Location/Qualifiers.
source
1. 2368
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 3.48e-286 Length: 2368
Score: 3936.00 Matches: 738
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-049-957-4 (1-738) x BD093599 (1-2368)
QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20
Db 61 ATGCGGGGTCGAGCGGGGCTCTGTGGCTGTCTGCGCTCTGCGCACCGTCTCGGAGGC 120
QY 21 MetGluValArgTyrCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40
Db 121 ATGAGGTGCGGTGGTGGCCACCTCGGACCCAGAGACGACCAAGTGGCGCAACATGAGC 180
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60
Db 181 GAGCGCTTCGCGGAAGCGGCGCATCCAGCCCTCCCTCTCTCGTCCGGGGCACCTCCGCC 240
QY 61 AspHisCysValGlnLeuLeuAlaAlaGlnAlaAspAlaIleThrLeuAspGlyGly 80
Db 241 GACCCTCGTCCAGCTCATTCGCGGCCGAGAGGCTGACGCGCATCTCTCGATGGAGGA 300
QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100
Db 301 GCCATCTATAGGCGGGAAAGAGACGCGCTGAAGCGGTGGTGGCGCAAGTGTACGAT 360
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValValArgSerSerHisValThr 120
```

Db 361 CAAGAGGTTCGGTACCTCTTATTACGCCCGTGGCTGTGGTCAGGAGAGCTCCCATGTGACC 420  
Qy 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140  
Db 421 ATTGACACCTGAAAGCGTGAAGTCTCTGCCACACGGGCATCAATCGACAGTGGGCTGG 480  
Qy 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
Db 481 AACGTGCCCGTGGCTTACCTGGTGGAGAGCGCGCTCTCGTGTATGGGCTCGCATGTA 540  
Qy 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyValAlaGlyGlnThrSer 180  
Db 541 CTCAAAGCTGTTCAGCGACTATTTTGGGGCAGCTGGTCCCGGGGCGAGAGAGACCATGT 600  
Qy 181 TyrSerGluSerLeuLysCysArgLeuLysCysArgGlyAspSerSerGlyGlyValCysAsp 200  
Db 601 TACTCTGAGTCTCTGTCTGCTCTGTCAGGGGTGACAGCTCTGGGGAGGGGTGTGTGAC 660  
Qy 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGlyGly 220  
Db 661 AAGAGCGCCCTCGAGAGATACAGACTACAGCGGGCCCTTCGGTGTCCCTGGCGAAGGG 720  
Qy 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuLysAsnThrAspGlyLysThr 240  
Db 721 GCAGGGGACGTGGCTTTTGTGAAGCACAGCAGCGTACTGGAGAACACCGGATGGGAAGCG 780  
Qy 241 LeuProSerTyrGlyGlnAlaLeuLysSerGlnAspPheGluLeuLysCysArgAspGly 260  
Db 781 CTTCCCTCTGGGGCCAGGCCCTGTGTCTACAGGACTTCGAGCTGTGTGCCGGATGGT 840  
Qy 261 SerArgAlaAspValThrGluTyrArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
Db 841 AGCGGGCGCATGTCAACGAGTGGAGCGAGTGCATCTGGCCGGGTGCTGCTCACGCC 900  
Qy 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLysGlnGlyGln 300  
Db 901 GTGGTGGTCCGGGCCACACAGATGGGGCCCTCATCTTCCGGTGTCTCAACGAAGGCCAG 960  
Qy 301 ArgLeuPheSerHisGlyGlySerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
Db 961 CGTCTGTTCAAGCACAGAGGGCCAGCAGCTTCCAGATGTTCCAGCTCTGAGGCCCTATGGCCAG 1020  
Qy 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
Db 1021 AAGGATCTACTTCAAGACTCTACTCTGGAGCTTGTGCCCATCGCCACACAGACCTAT 1080  
Qy 341 GluAlaTyrLeuGlyHisGlyTyrLeuHisAlaMetLysGlyLeuLysCysAspProAsn 360  
Db 1081 GAGGCGTGGTGGGCCCATGAGTACCTGCACGCCCATGAAGGGTCTGCTCTGACCCCAAC 1140  
Qy 361 ArgLeuProTyrLeuArgTyrCysValLeuSerThrProGluIleGlnLysCysGly 380  
Db 1141 CGGCTGCCCCCTTACCTGGCTGGTGTGTCTCTCCACTCCCGAGATCCAGAGTGTGA 1200  
Qy 381 AspMetAlaValAlaPheArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
Db 1201 GACATGGCGTGGCTTCCCGCGGACAGCGCTCAAGCCAGAGATCCAGTGGTGTGAGCC 1260  
Qy 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
Db 1261 AAGTCCCCCAACACTGCAATGGAGCGGATCCAGGCTGAGCAGGTGACGCTGTGACCCCTA 1320  
Qy 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
Db 1321 AGTGGCGAGACATTTACCGCGGGGAGAGATACGGCTGGTTCCTCCGAGCGCGGAG 1380  
Qy 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp 460  
Db 1381 CACTATGCCCGGAGACAGACAGCAACTCGTACTAGTGTGGCCGTGTGTGAGCGGAC 1440  
Qy 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
Db 1441 AGCTCCACGCGCTTCACTTGGATGAGTTCGGGGCAAGCGCTCTCTGCCACGCGGTTTC 1500

Qy 481 GlySerProAlaGlyTyrAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
Db 1501 GGCAGCCCTGAGCGGTGGGATGTCCTCCGTGGGTGCCCTTATTTCAGAGAGCTTCATCCGG 1560  
Qy 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
Db 1561 CCCAAGGACTGTGAGCTCTCACAGCAGTACGAGAGTTCCTTCAATGCCAGCTGCGGTGCC 1620  
Qy 521 ValAsnAsnProLysAsnTyrProSerSerLeuLysCysAlaLeuLysValGlyAspGluGln 540  
Db 1621 GTGAACAACCCCAAGAACTACCCCTCTCGCTGTGTGTGCACTGTGTGGGGAGCGAGCAG 1680  
Qy 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
Db 1681 GGCCTCCCAAGTGTGTGGCAACAGCCAGGAGCGGTATTACGGCTACCGCGCGCGCTTC 1740  
Qy 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
Db 1741 AGGTGCTGTGGAGAAATCGGGTGACGTTCCTTCGTGAGGCACACAACTCTCTTTGAC 1800  
Qy 581 AsnThrAsnGlyHisAsnSerGluProTyrAlaAlaGluLeuArgSerGluAspTyrGlu 600  
Db 1801 AACAAACGGCCCAATTCGAGCCCTGGGCTGTGAGCTCAGGTGAGAGACTATGAA 1860  
Qy 601 LeuLysCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
Db 1861 CTGCTGTGCCCAACCGGGCCGAGCGGAGGTGTCCAGTTTTCAGGCTGCAACCTGGCA 1920  
Qy 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
Db 1921 CAGATACACCCCGCCGCGGTATGTCGCGGCCGACACCAATCTTCACTCCGTGTATGA 1980  
Qy 641 LeuLysAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660  
Db 1981 CTGCTGGACAGGCCCGAGGACCTGTTGGAGACGACCAATAGAACGGGTTCAAAATG 2040  
Qy 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680  
Db 2041 TTCGACTCTCCAACTATCATGCGCAAGACCTGCTTTTCAAGGATGCCACCTCGGGCG 2100  
Qy 681 ValProValGlyGlyLysThrThrTyrArgGlyTyrLeuLysLeuAspTyrValAlaAla 700  
Db 2101 GTGCTGTGGAGAGAAAACCACTACCGCGGCTGGCTGGGCTGGACTACGTGGCGCG 2160  
Qy 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720  
Db 2161 CTGGAGGAGATGTGTCTCAGCAGTGTCTGGCGCGAGCGGCCCGCGGCGCGCGGCC 2220  
Qy 721 LeuLeuProLeuLeuLeuProAlaLeuAlaLysArgLeuLeuProProAlaLeu 738  
Db 2221 CTGCTCCCGTGTCTGCTGCGCGCCCTCGCGCGCGCTGTCTCCCGCGCGCTC 2274

## RESULT 7

BD107341  
LOCUS 2368 bp  
DEFINITION Novel chondrogenesis promoter.  
ACCESSION BD107341  
VERSION BD107341.1 GI:23202159  
KEYWORDS JP 20020311-A/4.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2368)  
AUTHORS Kato, Y. and Fujimoto, K.  
TITLE Novel chondrogenesis promoter  
JOURNAL Patent: JP 20020311-A 4 23-JAN-2002;  
JAPAN SCIENCE AND TECHNOLOGY CORP  
COMMENT OS Homo sapiens (human)  
PN JP 20020311-A/4  
PD 23-JAN-2002  
PF 07-JUL-2000 JP 2000206566

BD107341  
LOCUS 2368 bp  
DEFINITION Novel chondrogenesis promoter.  
ACCESSION BD107341  
VERSION BD107341.1 GI:23202159  
KEYWORDS JP 20020311-A/4.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2368)  
AUTHORS Kato, Y. and Fujimoto, K.  
TITLE Novel chondrogenesis promoter  
JOURNAL Patent: JP 20020311-A 4 23-JAN-2002;  
JAPAN SCIENCE AND TECHNOLOGY CORP  
COMMENT OS Homo sapiens (human)  
PN JP 20020311-A/4  
PD 23-JAN-2002  
PF 07-JUL-2000 JP 2000206566







Qy	221	AlaGlyAspValAlaAlaPheValIyHisSerThrValLeuGluAenThrAspGlyLysThr	240
Db	721	GCAGGGACGCTGGCTTTTGTGAAGCACAGCACGCTACTGGAGAACACGATGGGAAGACG	780
Qy	241	LeuProSerThrGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly	260
Db	781	CTTCCCTCTCGGGGCCAGGCCCTCTGTGCACAGACTTCGAGCTGTGTGCGGGATGGT	840
Qy	261	SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla	280
Db	841	AGCCGGGCCGATGTACCCAGTGGAGCAGTGCCATCTGGCCCCGGGTGCTGTCTACGCC	900
Qy	281	ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln	300
Db	901	GTGGTGGTCCGGGCCGACACAGATGGGGCCCTCATCTCCGGCTGCTCAACGAAGGCCAG	960
Qy	301	ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrglyGln	320
Db	961	CGTCTGTTCAGCCACGAGGGCAGAGCTTCCAGATGTTTCAGCTCTGAGGCTATGGCCAG	1020
Qy	321	LysAspLeuLeuPheLeuAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr	340
Db	1021	AAGGATCTACTCTTCAAGAAGCTCTACCTCGGAGCTTGTGCCATCTGCCACACAGACCTAT	1080
Qy	341	GluAlaTrpLeuGlyHisGluTyLeuHisAlaMetLysGlyLeuLeuCysAspProAsn	360
Db	1081	GAGCGTGGCTGGGCCCATGAGTACTGCACGCCATGAAGGGTCTGTCTGTGACCCCCAAC	1140
Qy	361	ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly	380
Db	1141	CGGTGCCCCCTACCTCGCTGGTGTGTCTCTCCACTCCGAGATCCAGAAGTGTGGA	1200
Qy	381	AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla	400
Db	1201	GACATGGCCGTGGGCTTCCGCCGCGAGCGCTCAAGCCAGAGATCCAGTGTGTGAGCC	1260
Qy	401	LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu	420
Db	1261	AAGTCCCCCAACACTCATGGAGCGGATCCAGGCTGAGCAGGTCTGACGCTGTGACCCYA	1320
Qy	421	SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu	440
Db	1321	AGTGGCAGGACATTTACCGCGGGGAAGAAGTACGGCTTGGTTCCTCCGAGCGCGCGAG	1380
Qy	441	HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp	460
Db	1381	CACTATGCCCGGAAGACAGCAGCACTCGTACTACGTGTGGCCGTGGTGTGAGACGGGAC	1440
Qy	461	SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe	480
Db	1441	AGTCTCCACGCCTTCACTTGGATGACTTCGGGGCAAGCGCTCTCTGCCACGCGGTTC	1500
Qy	481	GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg	500
Db	1501	GGCAGCCCTGCAGCTGGGATGTCCCGTGGGTGCCCTTATTTCAGAGAGCTTCATCCGG	1560
Qy	501	ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro	520
Db	1561	CCCAGAAGACTGTGACGTCTCACAGCAGTGAAGGAGTTCCTTCAATGGCAGCTGCTGCC	1620
Qy	521	ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln	540
Db	1621	GTGNACAAACCCCAAGAACTACCCCTCTCTGCTGTGTGCATCTGTGCGTGGGGGACGAG	1680
Qy	541	GlyArgAsnLysCysValGlyAsnSerGlnIleArgTyrTyrGlyTyrArgGlyAlaPhe	560
Db	1681	GGCCGCAACAAGTGTGTGGGCAACAGCCAGAGCGGTATTACGGCTACCGCGGCGCTTC	1740
Qy	561	ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp	580
Db	1741	AGGTGCTGTGTGGAAATGCGGGTGACGTTGCCCTTGTGTGAGGCACACCAACCGTCTTGAC	1800

Qy	581	AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu	600
Db	1801	AACACAAAGCGCCACAATTCGAGCCCTGGCTGCTGAGCTCAGGTCAAGGCAGAGGACTATGAA	1860
Qy	601	LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla	620
Db	1861	CTGCTGTGCCCCAACCGGGGCCGAGCGAGGTGTCCAGTTTCAGCCTGCACACCTGGCA	1920
Qy	621	GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly	640
Db	1921	CAGATACACCCACGCCGCTGATGTTCGGCCCGCACCAACAATCTTTCACCGCTGATGGA	1980
Qy	641	LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet	660
Db	1981	CTGCTGGACCAAGGCCCAAGGACCTGTTTGGAGACGACCACAAATAAGAACGGTTCAAAATG	2040
Qy	661	PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla	680
Db	2041	TTCGACTCTCCCACTATCATGGCCAAAGACCTGCTTTTCAAGGATGCCACCGTCCGGCGC	2100
Qy	681	ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyValLeuAspTyrValAlaAla	700
Db	2101	GTGCCTGTCCGAGAGAAAACACTACCGCGGCTGGCTGGGGCTGGACTACGTGGCGCGC	2160
Qy	701	LeuGluGlyMetSerSerGlnGlnCysSerGlyValAlaAlaProAlaProGlyAlaPro	720
Db	2161	CTGGAAGGGATGCTCTCAGCAGTGTCTGGGCGCAGCGGCCCGGGCGCGGCC	2220
Qy	721	LeuLeuProLeuLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu	738
Db	2221	CTGCTCCCGTGTGCTGTCCCGCCCTCGCCCGCGCTGTCTCCCGCCCGCCCTC	2274
RESULT	9		
LOCUS	BD012530	2154 bp DNA linear	PAT 02-AUG-2002
DEFINITION	Method for determining arthritis related melonotransferrin and a reagent thereof.		
ACCESSION	BD012530		
VERSION	BD012530.1	GI:22092719	
KEYWORDS	WO 0111368-A/1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Kato,Y., Matsukawa,H., Yoshiwara,Y., Oka,O. and Fujita,T.		
TITLE	Method for determining arthritis related melonotransferrin and a reagent thereof		
JOURNAL	Patent: WO 0111368-A 1 15-FEB-2001; ORIENTAL YEAST CO LTD, YUKIO KATO, HIROKAZU MATSUKAWA, YUTAKA YOSHIWARA, OSAMU OKA, TSUYOSHI FUJITA		
COMMENT	OS Homo sapiens (human) PN WO 0111368-A/1 PD 15-FEB-2001 PF 02-AUG-2000 WO 2000JP005171 PR 05-AUG-1999 JP 99p 222568 PI YUKIO KATO, HIROKAZU MATSUKAWA, YUTAKA YOSHIWARA, OSAMU OKA, PI TSUYOSHI FUJITA PC G01N33/574 CC		
FEATURES	Key	Location/Qualifiers.	
source	1..2154 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"		
ORIGIN			
Alignment Scores:	1.43e-278	Length:	2154
Pred. No.:	3834.00	Matches:	718
Score:		Conservatives:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%		



Query Match:	97.41%	Indels:	0
DB:	6	Gaps:	0
US-10-049-957-4 (1-738) x BD012530 (1-2154)			
QY	21	MetGluValArgTyrCysAlaThrSerAspProGluGlnHisCysGlyAsnMetSer	40
DB	1	ATGGAGGTGGTGGTGGCCACCTTCGGACCCAGAGCAGCAAGTGGCGCAACATGAGC	60
QY	41	GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla	60
DB	61	GAGGCTTCGGGGAAGCGGCATCCAGCCCTCCCTCTCGTCCGGGGCACCTCCGCC	120
QY	61	AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly	80
DB	121	GACCACCTGCTCAGCTCATCCGGCCAGAGGCTGAGCCCATCACTCTCGATGAGGA	180
QY	81	AlaIleTyrGluAlaGlyCysGlnHisGlyLeuValProValValGlyValTyrAsp	100
DB	181	GCCATCTATGAGCGGGAAGAGCAGCGCCCTGAAGCCGGTGGTGGCGAAGTGACGAT	240
QY	101	GlnGluValGlyThrSerTyrTyrAlaValAlaValValArgSerSerHisValThr	120
DB	241	CAAGAGTGGTACCTCTATTACGCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	300
QY	121	IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnAsnThrValGlyTyr	140
DB	301	ATTGACACCTGAAAGGCGTGAAGTCTGCCACACGGGCATCAATCGCACAGTGGGCTGG	360
QY	141	AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal	160
DB	361	AACGTGCCGTGGGTACCTGTGGAGAGCGCCGCTCTCGGTGATGGGCTGCCATGTA	420
QY	161	LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyLysThrSer	180
DB	421	CTCAAGCTGTGACGACTATTTTGGGGCAGCTGGTCCCGGGGCGAGGAGCAGT	480
QY	181	TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGlyValCysAsp	200
DB	481	TACTCTGAGTCCCTGTGCTGCTGACGGGTGACAGCTCTGGGGAGGGGTGTGTGAC	540
QY	201	LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGly	220
DB	541	AAGAGCCCTCGAGAGATACACACTACGCGGGCCCTTCGGTGGCTGGCGAAGGG	600
QY	221	AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr	240
DB	601	GCAGGGGACGTGGCTTTTGTGAAGCACACACGCTACTGGAGAACACGGATGGGAGACG	660
QY	241	LeuProSerTyrGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly	260
DB	661	CTTCCCTCTGGGGCCAGGCCCTGCTGTACAGGACTTCGAGCTGTGTGGCGGATGGT	720
QY	261	SerArgAlaAspValThrGluThrArgGlnCysHisLeuAlaArgValProAlaHisAla	280
DB	721	AGCCGGGCGCATGTACCCAGTGGAGGAGTCCATCTGGCCCGGTGCTCTGCTCAGGCC	780
QY	281	ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln	300
DB	781	GTGGTGGTCCGGGCCGACACATATGGGGCCCTCATCTCCGGCTGCTCAACGAAGGCCAG	840
QY	301	ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln	320
DB	841	CGTCTGTTCCAGCCACGAGGGCAGCAGCTTCCAGATGTTCCAGCTCTGAGGCCCTATGGCCAG	900
QY	321	LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr	340
DB	901	AAGGATCTACTCTTCAAGACTCTACCTCGGAGCTTGTGCCCATCGCCACACAGACCTAT	960
QY	341	GluAlaTyrLeuGlyHisGlnTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn	360
DB	961	GAGGCGTGGCTGGGCCATGAGTACCTGACCGCCATGAAGGGGTCTGCTCTGTGACCCCAAC	1020
QY	361	ArgLeuProProTyrLeuArgTyrCysValLeuSerThrProGluIleGlnLysCysGly	380
DB	1021	CGGCTGCCCCCTTACCTCGCTGGTGTGTCTCTCCACTCCCGAGATCCAGAAAGTGTGA	1080
QY	381	AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla	400
DB	1081	GACATGGCGGTGGCCCTTCGGCGGCGAGCGCTCAAGCCAGAGATCCAGTGGTGTGAGCC	1140
QY	401	LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu	420
DB	1141	AAGTCCCCCAACACTGTCATGAGCGGATCCAGGCTGAGCAGCTCGACGCTGTGACCTTA	1200
QY	421	SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu	440
DB	1201	AGTGGCGAGGACATTTACACGCGGGGAGAGTACGGCTGTGGTTCGCCGACCGCGGAG	1260
QY	441	HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp	460
DB	1261	CACATATGCCCGGAAGACAGCAGCAACTCGTACTACTGTGTGGTGGTGGTGGTGGTGG	1320
QY	461	SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe	480
DB	1321	AGTCCCCAGCCTTCACTTGGATGAGCTTCGGGGCAAGCGCTCTCTGCCACGCCGTTTC	1380
QY	481	GlySerProAlaGlyTyrAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg	500
DB	1381	GGCAGCCCTGAGGCTGGGATGTCCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1440
QY	501	ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro	520
DB	1441	CCCAAGGACTGTGAGCTCTCACAGCTGAGCGAGTCTTCTTCAATGCCAGCTTCGTGCC	1500
QY	521	ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln	540
DB	1501	GTGACAAACCCCAAGAACTACCCCTCTCGCTGTGTGCACTGTGCTGGGGACGAGCAG	1560
QY	541	GlyArgAsnLysCysValGlyAsnSerGlnLysArgTyrTyrGlyTyrArgGlyAlaPhe	560
DB	1561	GGCCCAACAAAGTGTGTGGCAACAGCAGGAGCGGTATTACGGCTACCGCGCGCCTTC	1620
QY	561	ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp	580
DB	1621	AGGTCCCTGGTGGAGAAATCGGGTGACGTTCCTTCGTGAGCACACAAACCTCTTTGAC	1680
QY	581	AsnThrAsnGlyHisAsnSerGluProTyrAlaAlaGluLeuArgSerGluAspTyrGlu	600
DB	1681	AACAAACGGCCACAATTCGAGCCCTGGGCTGTGAGCTCAGTCTCAGAGTATGAA	1740
QY	601	LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla	620
DB	1741	CTGCTGTGCCCAACCGGGCCGAGCCGAGGTGTCCAGTTCGAGCTTGCACACCTGGCA	1800
QY	621	GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly	640
DB	1801	CAGATACACCCCGCCGCGGTATGTCCGCGCCGACACCAACATCTTACCGTGTATGA	1860
QY	641	LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet	660
DB	1861	CTGCTGGACAGGCCCGCAGACCTGTTGGAGACGACCAATAAGACGGGTTCAAAATG	1920
QY	661	PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla	680
DB	1921	TTGACACTCTCAACTATCATGGCCCAAGACCTGTCTTTCAAGATGCGCACCTCCGGCG	1980
QY	681	ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla	700
DB	1981	GTGCTGTGGAGAGAAAACCACTTACCGCGCTGGTGGGCTGGACTTACCTGTGGCGCG	2040
QY	701	LeuGluGlyMetSerSerGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro	720
DB	2041	CTGGAGAGGATGTCTCTCAGCAGTGTCTGGGCGCAGCGGCCCGCGCGCGCGCGCC	2100
QY	721	LeuLeuProLeuLeuProAlaLeuAlaArgLeuLeuProProAlaLeu	738

[illegible]

```
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPheAsnAlaSerCysValPro 520
Db 1560 CCAGGAACCTGGACGCTCTCAAGCGGTGGGTCACTTCTCAACGCCAGCTGTGTGCGG 1619
QY 521 ValAsnAsnProLysAsnTyrProSerLeuLeuCysAlaLeuValGlyAspGluGln 540
Db 1620 GTGAACAACCCCAAGAGTACCCCTCTCGCTGTGGCACTCTGGGTGGGTGACGAGCAG 1679
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560
Db 1680 GGCCGCAACAGTCACTGGCAACAGCCAGGAGCGGTACTATGGCGACAGTGGCGCTTC 1739
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580
Db 1740 AGGTGCTGTGTGGAGGTTCAGGGAGCTGGGCTTTCGTCAAGCACACGACCATCTTTGAC 1799
QY 581 AsnThrAsnGlyHisAsnSerGluProTyrAlaAlaGluLeuArgSerGluAspTyrGlu 600
Db 1800 AACACAAATGGCCACAATCCCGAGCGGTGGGTGGCCATCTGAGGAGCCAGCACTACGAG 1859
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
Db 1860 CTGCTGTGCCCCAACCGCGCGAGCTGAGCGCACCATGTTGCCGCTGCACCTGGCC 1919
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
Db 1920 CAGATTCCGTCCAGCGCTCATGTGTGGCGCCGACCAACATCTTCAACCGTTTACGGA 1979
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660
Db 1980 CTGTGGGACAGCGCCAGGACCTGTTTGGAGACGACCAACAAAGAACGGGTTCAGATG 2039
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680
Db 2040 TTCGACTCTCCAGCTACACCGCGGAGACCTGCTCTTCAAGACGCGCACGGTGGCGCT 2099
QY 681 ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700
Db 2100 GTGCTGTGGCGGAGAGGACCACTTACCAGGACTGGCTGGGCGCGGACTACGCTGGCGGT 2159
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720
Db 2160 CTGGAAGGATGATCAGTCACAGCGTGTCTCAGGGGAGCGCTCGGCGCGCGCGGCGCTCG 2219
QY 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738
Db 2220 CTGCTGCGCGCTGCTG-----CCCTGGCTGGCGGCTCTCTGCTTTCGCTC 2267

RESULT 11
BD107340
LOCUS BD107340 2388 bp DNA linear PAT 18-SEP-2002
DEFINITION Novel chondrogenesis promoter.
ACCESSION BD107340
VERSION BD107340.1 GI:23202158
KEYWORDS JP 2002020311-A/3.
SOURCE Oryctolagus cuniculus (rabbit)
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
Kato, Y. and Fujimoto, K.
Novel chondrogenesis promoter
Patent: JP 2002020311-A 3 23-JAN-2002;
JAPAN SCIENCE AND TECHNOLOGY CORP
OS Oryctolagus cuniculus (rabbit)
PN JP 2002020311-A/3
PD 23-JAN-2002
PF 07-JUL-2000 JP 2000206566
PI YUKIO KATO, KATSUMI FUJIMOTO
PC A61K38/00, A61P19/00, A61P19/02, A61P19/08, C07K16/18, C12N15/09,
PC C12Q1/68,
PC G01N33/53, G01N33/566 //C07K14/42, C07K14/51, C07K14/79, A61K37/02,
```

```
PC C12N15/00
CC Novel chondrogenesis promoter
FH Key Location/Qualifiers
FT source 1..2388
FT /organism='Oryctolagus cuniculus (rabbit)'
FEATURES
source Location/Qualifiers
1..2388
/organism='Oryctolagus cuniculus'
/mol_type='genomic DNA'
/db_xref='taxon:9986'
```

## ORIGIN

## Alignment Scores:

```
Pred. No.: 5.53e-249 Length: 2388
Score: 3441.00 Matches: 640
Percent Similarity: 92.95% Conservative: 46
Best Local Similarity: 86.72% Mismatches: 50
Query Match: 87.42% Indels: 2
DB: 6 Gaps: 1
```

US-10-049-957-4 (1-738) x BD107340 (1-2388)

```
QY 1 MetArgGlyProSerGlyAlaLeuThrLeuLeuAlaLeuArgThrValLeuGlyGly 20
Db 60 ATGCGGTGCCGAAGCGCGCTATGTGGATCTTCTGGCCCTCGGCACCGCATCTCGGCAGC 119
QY 21 MetGluValArgTyrCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40
Db 120 GTGGAGGTGGGTGTGTGACCCGCTCCGAGCCCGAGCAGCAGCAAGTGGAGACATGAGC 179
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60
Db 180 CAGGCTTCCGGAAGCGCGCTCCAGCCCGCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGG 239
QY 61 AspHisCysValGlnLeuLeuAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80
Db 240 GACCACTCGCTCCAGCTCATCGGCGCCAGGAGCGCGCATCTCTGGACGGAGGA 299
QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100
Db 300 GCCATTACGAGCGGGGAGGAACACCGCTGAAGCCCGTGGTGGGCGGAGTGTATGAC 359
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValArgSerSerHisValThr 120
Db 360 CAAGAGGTGGGCACCTCTCTACTGTGCGCGTGGTCAAGAGGAGCTCCCAACGTGACC 419
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140
Db 420 ATCAACACCTTGAGAGCGGTGAAGTCTTGCCACACGCGGCATCAACCGCAGTGGGCTGG 479
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160
Db 480 AACGTGCTGTGGGTACTCTGGTGACAGCGCGCCCTCTCTAGTATGGGCTGTGACGTG 539
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyValAlaGlyThrSer 180
Db 540 CTCAAAGCGGTCAAGAGTACTTTCGGGGGAGCTCGCTCCCTGGGCGAGGAGACCA 599
QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200
Db 600 TACTCGGAGTCCCTCTGTGCTCTTCCGCGGGCGGACACCTCCGGGGAGGGGTGTGTGAC 659
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGlyGly 220
Db 660 AAGAGCCCTTGGAGCGGTACTACGACTTACAGCGGGGCTTCCGGTGGCTGGCAGAGGC 719
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240
Db 720 GCAGGGGACGTGGCCCTTTGTGAAGCACAGCAGCGGTGTGGAGAACACCGATGGGAGAAC 779
QY 241 LeuProSerTyrGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260
Db 780 CTGCCCCCTCTGGGGCCACATGCTGTATGTACGGGACTTTGAGCTGTGTGCGGGGACGCGC 839
```

```
QY 261 SerArgAlaAspValThrGluTirArgGlnCysHisLeuAlaAargValProAlaHisAla 280
Db 840 AGCCGGGCGAGGTCACAGTGGGAGACCTCCACCTGGCCCGGGTCCCGCCACGCC 899
QY 281 ValValAlaArgAlaAspThrAspGlyGlyLeuIlePheAargLeuLeuAasnGluGlyGln 300
Db 900 GTGGTGGTCCGGGCGGACACCGAGCGAGGCTCATCTTCCGGCTTCTCAATGAGGCGCAG 959
QY 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320
Db 960 CGGCTGTTCAGCCAGCGGCGAGCTTCAGATGTTTCAGTCTGGAGGCGCTTACGCCGAC 1019
QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340
Db 1020 AAGAACCCTGCTGTTCAAGACTCCACGCTGGAGCTGGTCCCATCGCCACACAGACCTAC 1079
QY 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360
Db 1080 GAGGCTGGTGGGCGGCGAGTACCTGCACGCCATGAAGGCTGTCTGTGACCCCAAC 1139
QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380
Db 1140 CGGCTGCCCGCCCATACCTCGCTGGTGGTGTCTCCACCCCGAGATCCAGAGGTGTGA 1199
QY 381 AspMetAlaValAlaPheArgGlnArgLeuLysProGluIleGlnCysValSerAla 400
Db 1200 GACATGGCGGTGGCCCTCAGCGGCGAGAGCTCAAGCCGGAGATCCAGTGTCTCGGCG 1259
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAlaAspValThrLeu 420
Db 1260 GAGTCCCCCGACCTGCTGATGAGCAGATCCAGGCTGGGCACATCGATGCTGTGACCCCTG 1319
QY 421 SerGlyGluAspIleTyrThrAlaGlyLysTyrGlyLeuValProAlaAlaGlyGlu 440
Db 1320 AACGGGAGGACATTCACACAGCGGGAGACTTATGGGCTGATCCCGGCTGCCGGGAG 1379
QY 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValAlaValArgAsp 460
Db 1380 CTGTATGCGGGGAGCAGCAGAGTACTCGTACTTCGGTGGGCGCGTGGTGAAGCAGAG 1439
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480
Db 1440 AGCGCTACGCTTCACCGTGGAGCAGCTGCGCGGAGAGCGCTCTGCCACCCCGGCTTC 1499
QY 481 GlySerProAlaGlyTyrAspValProValGlyAlaLeuIleGlnAargGlyPheIleArg 500
Db 1500 GGCAGCGCGCGGCTGGGAGCTCCCGTGGGCGGCTCATCCACTGGGGCTATATCCGG 1559
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520
Db 1560 CCCAGGAAGTGGAGCTCTCACAGCGTGGGTAGTTCCTTCAAGCCAGCTGTGTGCGG 1619
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540
Db 1620 GTGAACACCCCAAGAGTACCCCTCTCGTGTGGCAGCTGCTGGTGGGTGAGGAGCAG 1679
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrAargGlyAlaPhe 560
Db 1680 GGCGCGCAACAAGTGCATCGCAACAGCAGGAGCGGTACTATGGCGCAGTGGCGGCTTC 1739
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580
Db 1740 AGGTGCTGTGGAGGGTGCAGGGAGCTGGGCTTCGTCAAGCAGACACCATCTTTGAC 1799
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuAargSerGluAspTyrGlu 600
Db 1800 AACCAATGCCACAATCCCGAGCGGTGGCTGCCCATCTGAGGAGCCAGGACTACGAG 1859
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
Db 1860 CTGCTGTGCCCAACGGCGCGAGCTGAGGCGCAGCAGCATGTTGGCGCTGCAACCTGGCC 1919
```

```
QY 621 GlnIleProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
Db 1920 CAGATTCGTCACCGCGCTCATGGTGGGCGCCGACACCAACATCTTCACGTTTACGGA 1979
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660
Db 1980 CTGCTGGACAAGGCCAGGACCTGTTTGGAGACGACCAACAAGAACGGGTTCAAGATG 2039
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla 680
Db 2040 TTCGACTCTCTCCAGCTACACGCGCGAGACCTGCTCTTCAAGGAGCCGACGCTGCGCT 2099
QY 681 ValProValGlyGluLysThrThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaAla 700
Db 2100 GTGCTGTGGCGGAGAGGACCACTTACGAGACTGGCTGGGCGCGGACTACTGTGGCGCT 2159
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGlyAlaPro 720
Db 2160 CTGGAAGGATGTCAGTCACAGCGTGTCTCAGGGGCGAGCGCTCGGCGGCCCGGCGCTCG 2219
QY 721 LeuLeuProLeuLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738
Db 2220 CTGCTGCGCTGCTG-----CCCTGGCTCGGGCGCTCTCTGCTGCTCTGCTC 2267
```

## RESULT 12

```
AB010995 LOCUS 2388 bp mRNA linear MAM 24-OCT-1998
DEFINITION Oryctolagus cuniculus mRNA for p97, complete cds.
ACCESSION AB010995
VERSION AB010995.1 GI:3786307
KEYWORDS membrane-bound transferrin-like protein; MTF; melanotransferrin;
p97.
SOURCE Oryctolagus cuniculus (rabbit)
ORGANISM Oryctolagus cuniculus
REFERENCE 1 (sites)
AUTHORS Kawamoto, T., Pan, H., Yan, W., Ishida, H., Usui, E., Oda, R.,
Nakamasu, K., Noshiro, M., Kawashima-Ohya, Y., Fujii, M., Shintani, H.,
Okada, Y. and Kato, Y.
TITLE Expression of membrane-bound transferrin-like protein p97 on the
cell surface of chondrocytes
JOURNAL Eur. J. Biochem. 256 (3), 503-509 (1998)
MEDLINE 98451505
PUBMED 9780225
REFERENCE 2 (bases 1 to 2388)
AUTHORS Kawamoto, T.
TITLE Direct Submission
JOURNAL Submitted (09-PEB-1998) Takeshi Kawamoto, Hiroshima University
School of Dentistry, Department of Biochemistry; 1-2-3, Kasumi,
Minami-ku, Hiroshima 734, Japan
(E-mail:tkawamo@ipc.hiroshima-u.ac.jp, Tel.81-82-257-5688,
Fax:81-82-257-5688)
```

## FEATURES

```
source
1. 2388
/organism="Oryctolagus cuniculus"
/mol_type="mRNA"
/sub_species="Japanese White"
/db_xref="taxon:9986"
/sex="male"
/cell_type="chondrocyte"
60..2270
/function="membrane-bound transferrin-like protein"
/notes="melanotransferrin (MTF)"
/codon_start=1
/product="p97"
/protein_id="BAA33956.1"
/db_xref="GI:3786308"
/translation="MRCRSAMWIFALRPTALGSVEVWCTASPEQOQKCDMSQAFR
EAGLPALLCVQGSADHCVOLIAHEADAITLDGAIYAGKEHGLKPVVGEVYDQE
VGTSYAVAVVKSSNTINTLRGSCHTGINTVGMNVPVGLVDSRGMVDCDV
LKAVSEYFGSCVPAGETYSLSCLRCRGTSGEGVCKSPLERYDYSGFRCFLA
EGAGDVAFVKHSTVLENTDRTPLPSWGHMLMSRDFELLRCRDSRSVTEWQHCLAR
```

## CDS

PAHAVVVRADTDAGLIPRLNEQRLFSHEGSSFQMFSSSEAYQKNNLFPKSTLELVP.  
 IATQTYEAWLGPYELHMKGLCDPNRLPPYLRCVLSTPEIKQKCDMAVAFSRQLK  
 PETQCVASBPQMEQIOAGHIDAVTLANGEDIHTAGTKTYGLIPAGELYAADRSNS  
 YFVAVVAKDSVAFTVDELGRKRSCHPGFSGPWVDPVAGALIHGWYIRPRNCVLT  
 AVQSPFNASCVPNNPKYPSSILCALCVDEGRNKCTGNSOERYVYDGSAPRCLVEG  
 AGDAVAFKHITTFIDNTNGNHPWAAHLRSQDYELLCPNGARAAHQFPAACNLAQIPS  
 HAVWRPDNTNIFVYGLDLKQDLFGDHNKGFPMFSDSSYHGRDLLLPKDATVRAVP  
 VGERITTYQDLWGLPDYVAALLEGWQSRCSGAAPGASLLPLPLAVGLULLSL"

## ORIGIN

## Alignment Scores:

Pred. No.: 11e-248 Length: 2388  
 Score: 3437.00 Matches: 639  
 Percent Similarity: 92.82% Conservative: 46  
 Best Local Similarity: 86.59% Mismatches: 51  
 Query Match: 87.32% Indels: 2  
 DB: 4 Gaps: 1

US-10-049-957-4 (1-738) x AB010995 (1-2388)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20  
 DB 60 ATCGCGTGGCGAGCGCGGTATGTGATCTTCTGGCCCTGGCCACCGCACTCGGCAGC 119  
 QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
 DB 120 GTGGAGGTGGGTGGTGTGACCGGTCTCGAGCCCGAGCAGCAAGTGGCAGGACATGAGC 179  
 QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
 DB 180 CAGGCTTTCGCGAAGCGCGCTTCCAGCCCGCCCTGCTGTGCGTGCGGCGACCTCGGCC 239  
 QY 61 AspHisCysValGlnLeuIleAlaGlnGlnAlaAspAlaIleThrLeuAspGlyGly 80  
 DB 240 GACCACTGGCTCAGCTCATCGCGGCCCGCCAGCGCGCCGCACTCATCTGGAGCGAGGA 299  
 QY 81 AlaIleTrpGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTrpAsp 100  
 DB 300 GCCATTTTACAGGCGGGGAGGAACACCGGCTGAAGCCGCTGGTGGCGCAAGTGTATGAC 359  
 QY 101 GlnGluValGlyThrSerTrpAlaValAlaValAlaValArgSerHisValThr 120  
 DB 360 CAAGAGGTGGCACTTCTACTACGCTGTGGCGGTGTCAAGAGGAGGCTCCAACTGAGC 419  
 QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp 140  
 DB 420 ATCAACACCTGAGAGCGGTGAAGTCTGCCACACCGGCATCAACCGCGTGGCTGG 479  
 QY 141 AsnValProValGlyTrpLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
 DB 480 AACGTGCTGTGGGCTACCTGTGGAGCAGCGCGCGCTCTCAGTGATGGGCTGTGACGTG 539  
 QY 161 LeuLysAlaValSerAspTrpPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180  
 DB 540 CTCAAGAGCGGTACGCGAGTACTTTCGGGGGCGAGTGGCTCCCTGGGCGAGAGACCA 599  
 QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200  
 DB 600 TACTCGAGTCCCTCTGTGCGCTCTCGCGGGGCGACACCTCCGGGAGGGGTGTGTGAC 559  
 QY 201 LysSerProLeuGluArgTrpTrpAspTrpSerGlyAlaPheArgCysLeuAlaGlyGly 220  
 DB 660 AAGAGCCCCCTGGAGCGGTACTACGACTACAGCGGGGCTTCCGGTGGCTGGCAGAAGGC 719  
 QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
 DB 720 GCAGGGGAGCGTGCCCTTTTGAAGACACAGCAGCGGTCTGGAGAACACGATGGAGAAC 779  
 QY 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260  
 DB 780 CTGCGCTCTCGGGGCCACATGCTGATGTACGGGACTTTGAGCTGTGTCGGCGGACGC 839  
 QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280

DB 840 AGCGCGGGCCAGCGTTCACCGAGTGGCAGCACTGCCACTGGCCGGTGGCCGCCACGCG 899  
 QY 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300  
 DB 900 GTGGTGGTTCGGGGCCGACACCGCAGCGAGCTTCATCTTCGGCTTCTCAATGAGGCCAG 959  
 QY 301 ArgLeuPheSerHisGluGlySerPheGlnMetPheSerSerGluAlaTrpGlyGln 320  
 DB 960 CGGCTGTTTCAGCCACAGGCGCAGCAGTTCAGATGTTTCAGCTCGGAGGCTACGCGCAG 1019  
 QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTrp 340  
 DB 1020 AAGAAGCTGCTGCTTCAAAGACTCCACGCTGGAGCTGGTCCCTCCACACACAGACTTAC 1079  
 QY 341 GluAlaTrpLeuGlyHisGluTrpLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360  
 DB 1080 GAGGCTGGCTGGGCCCCGAGTACCTGCACGCCATGAGGGTCTGCTCTGTGACCCCAAC 1139  
 QY 361 ArgLeuProProTrpTrpLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
 DB 1140 CGGCTGCCCCCATACCTCGCTGGTGGTGTCTCCACCCCGAGATCCAGAAGTGTGGA 1199  
 QY 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
 DB 1200 GACATGGCGCTGGCTTTCAGCGCGCAGAGGCTCAAGCCGAGATCCAGTGTGTCTCGGC 1259  
 QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluValAspAlaValThrLeu 420  
 DB 1260 GAGTCCCCCAGCAGCTGCATGAGCAGATCCAGGCTGGGCACATCGATGCTGTGACCTTG 1319  
 QY 421 SerGlyGluAspIleTrpAlaGlyLysLysTrpGlyLeuValProAlaAlaGlyGlu 440  
 DB 1320 AACGGGAGGACATTCACACAGCGGGAAGACTTATGGCTGATCCCGCTGCGGGGAG 1379  
 QY 441 HisTrpAlaProGluAspSerSerAsnSerTrpTrpValValAlaValValArgAsp 460  
 DB 1380 CTGTATCGCGCGAGCAGCAGGAGTAACTCGTACTTCTGTGTGGCGGTGTGAGCGAGAC 1439  
 QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
 DB 1440 AGCGCTACGCTTTCACGTGGACGAGCTGCGGGGAGCGCTCTCCACCTGGGCTACATCCG 1499  
 QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheLeArg 500  
 DB 1500 GCAGCGCGCGCGCTGGGAGCTGCGGTGGCGGCTCATCCACTGGGCTACATCCCG 1559  
 QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPheAsnAlaSerCysValPro 520  
 DB 1560 CCCAGGAAGTGGCGAGTCTCTCAGCGGTGGGTCAAGTCTTCAACGCCAGCTGTGTGCCG 1619  
 QY 521 ValAsnAsnProLysAsnTrpProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
 DB 1620 GTGAACACCCCAAGAGTACCCCTCTCGCTGTGGGCACTCTGCGTGGGTGACGAGCAG 1679  
 QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTrpTrpGlyTrpArgGlyAlaPhe 560  
 DB 1680 GCGCGCAACAACTGGCACTGGCAACAGCAGGAGCGGTACTATGGCGACAGTGGCGCTTC 1739  
 QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
 DB 1740 AGGTGCTGTGGTGGGAGGTGACGGGAGCTGGGCTTCTGTCGAGCAGCAGCACTCTTTGAC 1799  
 QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTrpGlu 600  
 DB 1800 AACACAAATGGCCACAATCCCGAGCGGTGGGTGGCCCTCTGAGGAGCAGGAGCACTACGAG 1859  
 QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
 DB 1860 CTGCTGTGCCCCCAACGGCGCGAGCAGAGCGGCACCACTTTGCGCGCTGCAACCTGGCC 1919  
 QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTrpGly 640



```
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480
|||.....|
Db 1444 AGTCTCTACTCTTCACTCCCTGGAGCAGTCTGGCGGCAAGCGTCTCTCCACCCCTACTTGT 1503
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500
|||.....|
Db 1504 GGCAGCCAGCGGCGTGGGAGGTCGCCATCGGCTCCCTCATCCAGCGGGGCTTCATCCCGG 1563
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPheAsnAlaSerCysValPro 520
|||.....|
Db 1564 CCCAAGGACTGTGATGTCCTCACAGCGGTGAGCGAGTCTTCAATCCAGCTGGTGCT 1623
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540
|||.....|
Db 1624 GTCAACAACCTAAGAACTACCTCCGCACTATGTGCGCTCTCGCTGGGAGACGAGAAG 1683
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560
|||.....|
Db 1684 GGCCTGCAACAATGTGTGGGAGCAGCCAGGAGAGATACCTACGGCTACAGCGGGGCTTTC 1743
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580
|||.....|
Db 1744 AGTGTCTTGTGGAGCATGACGGGAGCTGGCTTCTGTCAGCACACGACTGTCTTTGAG 1803
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGly 600
|||.....|
Db 1804 AACACAATGGTCACAATCTCGAGCTTGGCTTCTCACTCAGGTGGCAAGACTATGAA 1863
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
|||.....|
Db 1864 CTACTGTGCCCCAATCGGGCAGCGGTGAGGTAGACCATGTTCAAGCTTGCACACCTGGCA 1923
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
|||.....|
Db 1924 CAAATGCCATCCACACCTGTATGTCGTCCTCCAGACCAACATCTTCACTGTGTATGGA 1983
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660
|||.....|
Db 1984 CTTCTGGCAAGGCCCGAGACCTGTTTGGAGACGACCAATAACAAGACGTTTCCAAATG 2043
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680
|||.....|
Db 2044 TTGACTCTCTCCAAATATCACAGCCAGACCTGTTTCAAGATGCTACAGTCCGAGCG 2103
QY 681 ValProValGlyGlyLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700
|||.....|
Db 2104 GTGCCAGTCCGGGAGAAACCACTACCTGGACTGGCTGGTCTGCTGCTGCTGCTGCTG 2163
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGlyAlaPro 720
|||.....|
Db 2164 CTGGAGGGGATGTGTCTCAGCAGTGTCTCGGTGACAGGCGCGGTGAGCGAGTCCCC 2223
QY 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProAlaLeu 738
|||.....|
Db 2224 CTGCTGGCCCTGCTCTCTGTGACCCCTGGTGCAGGCTCTCTCTCGCTTCTC 2277

RESULT 14
LOCUS BC040347
DEFINITION Mus musculus antigen p97 (melanoma associated) identified by
IMAGE:3672916), complete cds.
ACCESSION BC040347
VERSION BC040347.1 GI:25955673
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 4102)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schetz,T.B., Brownstein,M.J., Udín,T.B., Toehiyuki,S.,
Carninci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,
Armsbron,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallu,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED 12477932
2 (bases 1 to 4102)
Strausberg,R.
Direct Submission
Submitted (27-NOV-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabps@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseg, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 23 Row: h Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13994139.
FEATURES
source
1..4102
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:18315 IMAGE:3672916"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI CGAP_Mam6"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
1..4102
/genes="Mf12"
/notes="synonyms: Mf1, CD228"
/db_xref="LocusID:30060"
/db_xref="MGI:1353421"
41..2257
/codon_start=1
/product="antigen p97 (melanoma associated) identified by
monoclonal antibodies 133.2 and 96.5"
/protein_id="AAH40347.1"
/db_xref="GI:25955674"
/db_xref="LocusID:30060"
/translation="NRLLSVTFWLLLSLRTVVVMEVQWCTISDAEQKCKDMSEAFQ"
```







```
QY 581 AsnThrAsnGlyHisAsnSerGluProTyrAlaAlaGluLeuArgSerGluAspTyrGlu 600
Db 1781 AACACAAATGGTCAACAATCTGAGCCTTGGCTTCTCACTCAGTGGCAAGACTATGAA 1840
QY 601 LeuLeuCyProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
Db 1841 CTACTGTGCCCCAATGGGGCAGCGGCTGAGGTAGACCAAGTTCACCACTTGGCAACCTGGCA 1900
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
Db 1901 CAAATGCCATCCACACGCTGCATGGTCCGTCCAGACACCAACATCTTCACTGTGTATGGA 1960
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660
Db 1961 CTTCTGGCAAGGCCAGGACCTGTTTGGAGACGACCAATAACAAGACGGTTTCCAAATG 2020
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla 680
Db 2021 TTTGACTCTCTCCAAATATCACAGCCAGACCTGCTTTTCAAAGATGCTACAGTCCGAGCG 2080
QY 681 ValProValGlyLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700
Db 2081 GTGCCAGTCCGGAGAAACCAACATACCTGGACTGGCTGGTCTGACTATGTGGTTGCG 2140
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGlyAlaPro 720
Db 2141 CTGGAGGGGATGTGTCTCAGCAGTGTCTCCGGTGCAGGGGCCGCGGTGCAGCGAGTCCC 2200
QY 721 LeuLeuProLeuLeuLeuProAlaLeuAlaAlaArgLeuLeuProAlaLeu 738
Db 2201 CTGCTGGCCCTGCTCTCTGTGACCCCTGGCTGCAGGCTCTCTCTCCGCTTCTC 2254
```

## RESULT 15

```
BD093609 4158 bp DNA linear PAT 27-AUG-2002
LOCUS Chondrogenesis promoter.
DEFINITION BD093609
ACCESSION BD093609.1 GI:22639197
VERSION WO 0113951-A/12.
KEYWORDS Mus sp.
SOURCE Mus sp.
ORGANISM Mus sp.
```

```
REFERENCE 1 (bases 1 to 4158)
AUTHORS Kato, Y. and Fujimoto, K.
TITLE Chondrogenesis promoter
JOURNAL Patent: WO 0113951-A 12 01-MAR-2001;
CHUGAI PHARMACEUTICAL CO LTD, YUKIO KATO, KATSUMI FUJIMOTO
COMMENT OS Mus sp. (mouse)
PN WO 0113951-A/12
PD 01-MAR-2001
```

```
PP 21-AUG-2000 WO 2000JP005590
PR 19-AUG-1999 JP 99P 232966
PI YUKIO KATO, KATSUMI FUJIMOTO
PC A61K45/00, A61K38/40, A61K48/00, A61K31/7088, A61K35/32, A61P19/02,
PC C07K14/47,
PC C07K14/79, C12Q1/02, G01N33/50, G01N33/15
CC
```

## FEATURES

```
source
1..4158
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 4,38e-238 Length: 4158
Score: 3300.00 Matches: 618
Percent Similarity: 89.84% Conservative: 45
Best Local Similarity: 83.74% Mismatches: 75
Query Match: 83.84% Indels: 0
DB: 6 Gaps: 0
```

```
US-10-049-957-4 (1-738) x BD093609 (1-4158)
QY 1 MetArgGlyProSerGlyAlaLeuTyrPLeuLeuAlaLeuArgThrValLeuGlyGly 20
Db 118 ATGAGGCTCTGAGCGGTGACTTTTGGCTACTCTCTGCTGCGCATCTGCTGTGTG 177
QY 21 MetGluValArgTyrCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40
Db 178 ATGGAGGTGACGTGGTGTACCATCTCAGACGCGAGCAGCAGAGAGTGCACACATGAGC 237
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60
Db 238 GAGGCGCTTCCAGGAGCTGGCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 297
QY 61 AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80
Db 298 GACCACTGTGTCCAGTCTCATCAAGAAACAAAGAGAGATGCCATCACTTGGATGAGGG 357
QY 81 AlalIeTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100
Db 358 GCCATCTATGAGCGAGGAGAGCAGCAGCGCTGAAGCCAGTGTGGGGAGTCTATGAC 417
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValAlaValArgSerSerHisValThr 120
Db 418 CAAGACATTTGGGACTTCTTATTATGCGGTGGTGTGTGTCAGGAGGAATTTCAATGTTAC 477
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140
Db 478 ATCAACACCTCAAGGGCGTCAAGTCTGCGCACACAGGCATTAACCGGACTGTGGGCTGG 537
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160
Db 538 AACGTGCGCTGTCTGTTTACCTCGTAGAGCGGCCATCTGTCTAGTGGGCTGTGATGG 597
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180
Db 598 CTCAAGCGCTGGTGGTGAATTTTGGAGGCGAGTGTGTCTCTGGAACAGGAGAAACAGC 657
QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyCyluGlyValCysAsp 200
Db 658 CATTCCGAGTCTCTCTGCGCTCTGCGTGGCGACTCTCTCTGCGGCACATGTGTGTGAC 717
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly 220
Db 718 AAGAGTCCCTTAGAGAGATACCTACGATACAGTGGAGCCTTCCGGTGGCTTGGCGAAGA 777
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240
Db 778 GCGCGTGCAGTGGCCTTCTGTAAGCACACAGCAGTGTGGAATACTGATGGAACACC 837
QY 241 LeuProSerTyrGlyGlnAlaLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260
Db 838 CTGCGCTTCTGGGGCAAGTCCCTGATGTCAGAGAGACTTCCAGCTACTATGACAGGGATGGC 897
QY 261 SerArgAlaAspValThrGluTyrArgGlnCysHisLeuAlaArgValProAlaHisAla 280
Db 898 AGCCGAGCGCATCATCTGAGTGGAGACGTTGCCACCTGGCCCAAGGTGCTGCTCATGCT 957
QY 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300
Db 958 GTGTGGTTCAGGGGTGACATGATGCGGTCTCATATTCCAATCTGCTCAACGAGGCCAG 1017
QY 301 ArgLeuPheSerHisGluGlySerPheGlnMetPheSerCyluAlaTyrGlyGln 320
Db 1018 CTTCTGTTCCAGCCACGAGACAGCTTCCAGATGTTCCAGTGTTCAGTCCAAAGGCCAG 1077
QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340
Db 1078 AAGAAGTCTGCTTCAAGACTTCCACTTGGAGCTTGTGCCCATTTGCCACACAGAACTAC 1137
QY 341 GluAlaTyrLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360
Db 1137 AAGAAGTCTGCTTCAAGACTTCCACTTGGAGCTTGTGCCCATTTGCCACACAGAACTAC 1137
```

Db 1138 GAGGCTGGCTGGCGCAGGAATACCTGCAGGCCATCAAGGGGCTCTCTGTGTATCCCAAC 1197  
Qy 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
Db 1198 CGGCTGCCCTACTACTCGCTGGTGTGTGTCTGAGCGCCGAGATCCAGAAGTGTGGA 1257  
Qy 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
Db 1258 GATATGGCTGTGGCTTTCAGCGCCGAGAAATCTCAAGCCGGAATTCAGTGTGTGTGCGCC 1317  
Qy 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluInValAspAlaValThrLeu 420  
Db 1318 GAGTCCCCTGAGCACTGCATGAGCAGATCCAGGCTGGGCACACTGACGCTGTGACTCTG 1377  
Qy 421 SerGlyGluAspIleTyrThrAlaGlyLysIleTyrGlyLeuValProAlaAlaGlyGlu 440  
Db 1378 AGGGCGCAGGACATTTACAGGGCAGGAAGGTGTACGGCCTGGTTCCGGCGCGCGGGAG 1437  
Qy 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp 460  
Db 1438 CTGTATGCTGAGGAGACAGGAGCAATTCCTACTTTGGTGGCTGTGGCAAGAGGGAC 1497  
Qy 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
Db 1498 AGCTCCTACTCCTTACCTGACGAGCTTCGCGCAAGCGTTCTGCCACCCCTACTTG 1557  
Qy 481 GlySerProAlaGlyTyrAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
Db 1558 GGCACGCCAGCGGGTGGAGGTGCCATCGCTCCCTCATCCAGCGGGCTTCATCCGG 1617  
Qy 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
Db 1618 CCCAAGGACTGTGATGTCTCACCGGTGAGCCAGTTCTTCATGCCAGCTGCGTGCCT 1677  
Qy 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluIn 540  
Db 1678 GTCAACAACCCCTAAGAACTACCTTCCGCACTATGTGCGCTCTGCGTGGGAGACGAGAAG 1737  
Qy 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
Db 1738 GCGCCCAACAATGTGTGGGAGCAGCGGAGAGATACCGGCTACAGCGGGGCTTC 1797  
Qy 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
Db 1798 AGTGTCTTGTGAGCATGCAGGAGACGTGGCTTTCGTCAGACACAGACTGTCTTTGAG 1857  
Qy 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600  
Db 1858 AACACAAATGGTCACAAATCCTGAGCCTTGGGCTTCTCACCTCAGGTGGCAAGACTATGAA 1917  
Qy 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
Db 1918 CTACTGTGCCCAATGTGGGCGACGGGCTGAGGTAGACCAGTTCCAAAGCTTGCAACCTGGCA 1977  
Qy 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
Db 1978 CAAATGCCATCCACGCTGTATGTTCGTCAGACACCAACATCTTCACTGTGTATGGA 2037  
Qy 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660  
Db 2038 CTTCTGGACAGGCCGAGGACCTGTTTGGAGACGACCAATAACAAGACGGTTTCCAAATG 2097  
Qy 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla 680  
Db 2098 TTTGACTCCTCCAAATATCACAGCAAGACCTGCTTTCAAAGATGCTACAGTCCGAGCG 2157  
Qy 681 ValProValGlyGlyLysThrThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaAla 700  
Db 2158 GTGCCAGTCCGGGAGAAACCACTACCTGGACTGGTGGTCTGACTATGTGTGGTGG 2217  
Qy 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGlyAlaPro 720  
Db 2218 CTGGAGGGGATGTTGTCTCAGCAGTGTCTCCGTTGCGGGGCGCGGTGCGAGTCCGCC 2277

Qy 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738  
Db 2278 CTGCTGGCCCTGCTCCTGCTGACCTGGCTGAGGCCTCCTTCCTCGCGTTCTC 2331

Search completed: May 16, 2004, 22:13:35  
Job time : 7793 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 16, 2004, 15:41:36 ; Search time 787 Seconds  
(without alignments)  
3983.698 Million cell updates/sec

Title: US-10-049-957-4

Perfect score: 3936

Sequence: 1 MRGPGCALWLLALRTVLGG.....APFLPLLLPALAARLLPAL 738

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-Q/cn2\_1/USPTO\_spool/US10049957/runat 14052004.100249.1247/app\_query.fasta\_1.903  
-DB=N\_Geneseq 29Jan04 -QFT=fastep -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10049957 @CGN 1.1 479 @runat 14052004.100249.1247 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq 29Jan04: \*  
1: Geneseqn1980s: \*  
2: Geneseqn1990s: \*  
3: Geneseqn2000s: \*  
4: Geneseqn2001as: \*  
5: Geneseqn2001bs: \*  
6: Geneseqn2002as: \*  
7: Geneseqn2003as: \*  
8: Geneseqn2003bs: \*  
9: Geneseqn2003cs: \*  
10: Geneseqn2004s: \*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3936	100.0	2368	2 AAF62197	Aaf62197 DNA encod
2	3936	100.0	2368	6 AAF62197	Abk12568 DNA encod
3	3936	100.0	2368	6 AAF62197	Aan70607 Sequence
4	3918	99.5	2362	1 AAN70607	Aaf63460 Human MTF
5	3834	97.4	2154	4 AAF63460	Aaf62196 DNA encod
6	3441	87.4	2388	6 AAF62196	Abk12568 DNA encod
7	3441	87.4	2388	6 AAF62196	Abk12568 DNA encod
8	3300	83.8	4068	4 AAD14466	Aad14466 Mouse p97

9	3300	83.8	4158	4 AAF62198	Aaf62198 DNA encod
10	1304.5	33.1	2347	2 AAF62198	Aaf62198 DNA encod
11	1272.5	32.3	2327	2 AAF62198	Aaf62198 DNA encod
12	1270.5	32.3	2318	7 AB282283	Abz34960 Human tra
13	1270.5	32.3	2327	3 AAZ24196	Aaz24196 Human ser
14	1270.5	32.3	2347	6 ABK84172	Abk84172 Human cDN
15	1270.5	32.3	2347	6 ABN96543	Abn96543 Gene #304
16	1270.5	32.3	2347	7 ACA64933	ACA64933 Human tra
17	1269.5	32.3	2097	7 AAV83200	AAV83200 Sequence
18	1268	32.2	2311	6 AB234960	Abz34960 Human gen
19	1265.5	32.2	4603	2 AAT45362	Aat45362 LDLR/TF c
20	1263.5	32.1	3428	2 AAT45361	Aat45361 LDLR/TF c
21	1263	32.1	2360	2 AAT48030	Aat48030 Human lac
22	1263	32.1	2360	2 AAV30770	AAV30770 Human lac
23	1263	32.1	2360	2 AAV80375	AAV80375 Human lac
24	1263	32.1	2360	4 AAA63788	AAa63788 cDNA enco
25	1263	32.1	2360	4 AAD06281	Aad06281 Human lac
26	1261	32.0	2117	2 AAO23453	Aao23453 Human lac
27	1261	32.0	2117	2 AAO29274	Aao29274 Human lac
28	1260.5	32.0	2124	2 AAO23454	Aao23454 Human lac
29	1260.5	32.0	2124	2 AAO29275	Aao29275 Human lac
30	1260.5	32.0	2136	6 ABA99732	ABA99732 Bovine BG
31	1258.5	32.0	2332	2 AAO06625	Aao06625 Human lac
32	1256.5	31.9	2358	2 AAO51346	Aao51346 Human lac
33	1256	31.9	2619	2 AAO54230	Aao54230 Human lac
34	1256	31.9	2619	2 AAV43729	AAV43729 Cancer as
35	1256	31.9	2619	6 ABL65858	ABl65858 Lung canc
36	1256	31.9	2619	6 AB235478	Abz35478 Human gen
37	1253.5	31.8	2319	2 AAO54229	Aao54229 Human lac
38	1252	31.8	2079	3 AAZ58122	Aaz58122 Synthetic
39	1252	31.8	2079	6 ABS66515	ABs66515 DNA enco
40	1252	31.8	2079	6 AAD45297	Aad45297 Human cod
41	1252	31.8	2079	6 AAD44922	Aad44922 Human cod
42	1252	31.8	2086	2 AAT02999	Aat02999 Lactoferr
43	1250.5	31.8	2347	2 AAT48031	Aat48031 Bovine la
44	1250.5	31.8	2347	2 AAV30771	AAV30771 Bovine la
45	1250.5	31.8	2347	2 AAV80376	AAV80376 Bovine la

ALIGNMENTS

RESULT 1	AAQ55768	AAQ55768 standard; cDNA; 2368 BP.
ID	AAQ55768	
XX	AC	AAQ55768;
XX	AC	AAQ55768;
DT	25-MAR-2003	(revised)
DT	20-JUL-1994	(first entry)
XX	XX	Human melanoma-associated antigen p97 coding sequence.
DE	DE	Human melanoma-associated antigen p97; Melanotransferrin; GPI;
KW	KW	glycosyl-phosphatidylinositol anchor; truncated, soluble;
KW	KW	blood-brain barrier; iron binding; Alzheimer's disease; iron metabolism;
KW	KW	disorder; haemochromatosis; ischaemic tissue damage; heart disease;
KW	KW	skin cancer; brain tumour; ss.
XX	XX	Synthetic.
OS	OS	
FH	Key	Location/Qualifiers
FT	CDS	61..2277
FT	FT	/*tag= a
FT	sig_peptide	61..117
FT	FT	/*tag= b
FT	mat_peptide	118..2274
FT	FT	/*tag= c
FT	FT	/product= "p97"
PN	WO9401463-A1.	
XX	XX	
PD	20-JAN-1994.	
XX	XX	

PF 09-JUL-1993; 93WO-CA000272.  
 XX PR  
 XX 10-JUL-1992; 92US-00912291.  
 XX (UYBR-) UNIV BRITISH COLUMBIA.  
 XX Jefferies WA, Moeger FL, Rothenberger S, Food MR, Yamada T;  
 XX PI  
 XX WPI: 1994-034993/04.  
 XX P-PSDB; AAR47899.  
 XX  
 PT New GPI-anchored and soluble forms of p97 - for treating disorders of  
 PT iron metabolism, delivering drugs to the brain and for diagnosis and  
 PT treatment of Alzheimer's disease.  
 XX  
 XX PS Disclosure; Page 102-106; 166pp; English.  
 XX  
 CC The melanoma-associated antigen p97 has been found to be a GPI-anchored  
 CC protein expressed on the cell surface which is able to bind iron. A  
 CC soluble form of p97 can be isolated from the aqueous phase after Triton-X  
 CC -114 phase separation. The soluble (hydrophilic) form of p97 does not  
 CC contain ethanolamine and it has a slower rate of transport than GPI-  
 CC anchored p97. The role of p97 in iron-transport suggests a use in  
 CC modulating iron uptake by cells; p97, its agonists, antagonists and  
 CC stimulants may be useful in treatment of conditions where iron-metabolism  
 CC is disturbed, e.g. haemochromatosis and skin cancer. Expression of p97  
 CC (and transferrin receptors) by reactive microglial cells associated with  
 CC senile plaques in Alzheimer's Disease (AD) provides a means of diagnosing  
 CC AD. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 2368 BP; 427 A; 766 C; 769 G; 406 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0 Length: 2368  
 Score: 3936.00 Matches: 738  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0

US-10-049-957-4 (1-738) x AAQ55768 (1-2368)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20  
 DB 61 ATGCGGGTCCGAGCGGGGCTCTGTGGTCTCTGGCTCTGCGCACCGTCTCGAGGC 120  
 QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
 DB 121 ATGAGGTTGGTGGTGGCCACCTCGGACCCAGAGCAGCAAGTGGCGCAACATGAGC 180  
 QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
 DB 181 GAGGCGTTCCGGGAACGGGCAATCCAGCCCTCCCTCTGCGTCCGGGGCACCTCCGCC 240  
 QY 61 AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80  
 DB 241 GACCACTGGTCCACTATCCGCGCCAGGAGGCTGACCGCATCACTCTGGATGGAGA 300  
 QY 81 AlaIleTrpGluAlaGlyLysGluHisGlyLeuLysProValValGlyValTrpAsp 100  
 DB 301 GCCATCTATAGCGGGGAAGAGGACACGCGCTGAGAGCGGTGGTGGCGAAGTGTACGAT 360  
 QY 101 GlnGluValGlyThrSerTrpTrpAlaValAlaValArgArgSerHisValThr 120  
 DB 361 CAAGAGGTCGGTACCTCTATTACCGCGTGGTGTGGTTCAGGAGGAGCTCCCATGTGACC 420  
 QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp 140  
 DB 421 ATTGACACCTTGAAGAGGCTGAGTCTTCCACACGCGGCATCAATCGCACAGTGGGCTGG 480  
 QY 141 AsnValProValGlyTrpLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
 DB 481 AACGTGCCGTTGGGTACTCTGGTGAGAGCGCGCCCTCTCGGTGATGGGCTCGCATGTA 540

QY 161 LeuLysAlaValSerAspTrpPheGlyGlySerCysValProGlyValGlyValThrSer 180  
 DB 541 CTCAAGCTGTACGAGCTATTTTGGGGGAGCTGCGTCCCGGGGAGGAGACCACTAGT 600  
 QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGlyValCysAsp 200  
 DB 601 TACTCTGAGTCCCTCTGTCCGCTCTGCAGAGGGTGACAGCTCTGGGAAGGGGTGTGTGAC 660  
 QY 201 LysSerProLeuGluArgTrpTrpAspTrpSerGlyAlaPheArgCysLeuAlaGlyGly 220  
 DB 661 AAGAGCCCTCTGGAGAGATACAGACTACAGCGGGGCTCTCCGTGCTCTGGCGGAAGGG 720  
 QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
 DB 721 GCAGGGGACGTGGGCTTTTGTGAAGCACAGACGCTACTGGAGAACACCGATGGGAAGACG 780  
 QY 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260  
 DB 781 CTTCCCTCTCGGGCCAGGCCCTGCTGTCAAGAGCTTCGAGCTGCTGTGCGGGATGGT 840  
 QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
 DB 841 AGCGGGGCGATGTCAACGAGTGGAGGCGAGTCCATCTGGCCCGGCTGCTGCTCACGCC 900  
 QY 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300  
 DB 901 GTGTGGTTCGGGGCGGACACAGATGGGGGCTCATCTTTCCGGCTGCTCAACGAAGGCCAG 960  
 QY 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTrpGlyGln 320  
 DB 961 CGTCTGTTCAGCCAGGAGGCGAGCTTCAGATGTTTCAGCTCTGAGGCCCTATGGCCAG 1020  
 QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTrp 340  
 DB 1021 AAGGATCTACTCTTCAAGAGCTACTACCTCGGAGCTTGTGCCATCGCCACACAGACCTAT 1080  
 QY 341 GluAlaTrpLeuGlyHisGluTrpLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360  
 DB 1081 GAGCGTGGCTGGGCGCATGAGTACTCGACGCCCATGAAGGGTCTGCTCTGTGACCCCAAC 1140  
 QY 361 ArgLeuProTrpTrpLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
 DB 1141 CGGCTGCCCCCTACTCTGGCTGGTGTGTCTCTCCACTCCCGAGATCCAGAGTGTGA 1200  
 QY 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
 DB 1201 GACATGGCGTGGCTTCCGCGCGAGCGCTCAAGCCAGAGATCCAGTGGCTGTGACGC 1260  
 QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
 DB 1261 AAGTCCCCCAACACTGCATGAGCGGATCCAGGCTGAGCAGGTGACGCTGTGACCCCTA 1320  
 QY 421 SerGlyGluAspIleTrpThrAlaGlyLysLysTrpGlyLeuValProAlaAlaGlyGlu 440  
 DB 1321 AGTGGGAGGACATTTACCGCGGGGAGAGAGTACGCGCTTGGTTCGCGAGCGCGGAG 1380  
 QY 441 HisTrpAlaProGluAspSerSerAsnSerTrpTrpValValAlaValAlaArgArgAsp 460  
 DB 1381 CACTATGCCCGGAAGACAGCAGCAACTCGTACTAGTGGTGGCGTGGTGGTGGAGCGGAC 1440  
 QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
 DB 1441 AGCTCCCAAGCTTCACTTGGATGAGCTTCGGGGGAGCGCTCTCGCACCGCGGTTTC 1500  
 QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
 DB 1501 GGCAGGCTCGAGCTGGGATGTCCCGTGGGTGCCCTTATTTCAGAGAGGCTTTCATCCGG 1560  
 QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
 DB 1561 CCCAAGGACTGTGAGCTCTCACAGCAGTGGAGCGGATTTCTTCAATGCCAGCTGCGTGCCC 1620



721	Db	GCAGGGGACGTGGCTTTTGTGAGACACAGACCGTACTCGAGAAACACGGATGGGAAGACG	780
241	Qy	LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGlnLeuLeuLeuCysArgAspGly	260
781	Db	CTTCCCTCCCTGGGGCCAGGCCCTGCTGTACAGGACTTCGAGCTGCTGTGCCGGGATGGT	840
261	Qy	SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla	280
841	Db	AGCCGGGCCGATGTACCGAGTGGAGGACGTGCCATCTGCCCCGGGTGCTGTCTACGCC	900
281	Qy	ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgIleuLeuAsnGluGlyGln	300
901	Db	GTGTGTGTCGGGGCCGACACAGATGGGGGCTCATCTTCGGCTGTCTCAACGAAGGCCAG	960
301	Qy	ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln	320
961	Db	CGTCTGTGTAGCCACAGAGGCAGAGCTTCCAGATGTTCAAGCTCTGAGGCTATGGCCAG	1020
321	Qy	LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr	340
1021	Db	AAGGATCTACTTTCANAAGACTCTACCTCGAGCTTGTGCCCATCGCCACACAGACCTAT	1080
341	Qy	GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn	360
1081	Db	GAGCGGTGGCTGGGCCATGATGACTTCGACGCCATGAAGGCTGTGCTCTGTGACCCCAAC	1140
361	Qy	ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly	380
1141	Db	CGGCTGCCCCCTACCTGCGCTGGTGTGCTCTCCACTCCGAGATCCAGAAGTGTGA	1200
381	Qy	AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla	400
1201	Db	GACATGGCCGTGGCTTCCGCGGCAGCGCTCAAGCCAGAGATCCAGTGGTGTGAGCC	1260
401	Qy	LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu	420
1261	Db	AAGTCCCCCAACACTGCATGGAGCGGATCCAGCTGAGCAGGTTCGACGCTGTGACCCTA	1320
421	Qy	SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu	440
1321	Db	AGTGGCGAGGACATTTACACGGCGGGGAAGATCGCCCTGGTTCGCGAGCGCGCAG	1380
441	Qy	HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValArgArgAsp	460
1381	Db	CACTATGCCCGGNAGACAGACGACNACTCTGACTACGTGTGGCCGTGTGTGACGCGGAC	1440
461	Qy	SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe	480
1441	Db	AGCTCCACACGCTTCACCTTGGATGAGCTTCGGGCAAGCGCTCTGCCACGCGGTTTC	1500
481	Qy	GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg	500
1501	Db	GGCAGCCCTCGAGCTGGGATGTCGCCGTGGGTGCCCTTATTACAGAGAGCTTCATCCGG	1560
501	Qy	ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro	520
1561	Db	CCCAAGGACTGTGACGTCTTCACAGCATGTAGCGAGTTCTTCAATGCCAGCTCGGTGCC	1620
521	Qy	ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln	540
1621	Db	GTGAACAACCCCAAGAACTACCCCTCTCGCTGTGTGCACTGTGTGGGGGACGAGCAG	1680
541	Qy	GlyArgAsnLysCysValGlyValAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe	560
1681	Db	GGCCGCAACAAGTGTGTGGGCACACAGCAGAGCGGTATTACGGCTACCGCGCGCGCTTC	1740
561	Qy	ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp	580
1741	Db	AGTGTCCCTGTGAGAAATGCGGTGACGTTGCTTGTGAGGCACACACCGCTTTTGAC	1800
581	Qy	AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu	600

Db	1801	AAACACAAACGGCCACAAATTCGAGCGCCCTGGCTGCTGAGCTCAGGTGACAGAGGACTATGAA	1861
Qy	601	LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla	620
Db	1861	CTGCTGTGCCCCACCGGGCCCGAGCCGAGGTGTCCAGTTTCAGCCTTGCAACCTGGCA	1920
Qy	621	GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly	640
Db	1921	CAGATACCAACCCACCGCGTGTGTCGCGCCCGACCAACATCTTCACCGTGTATGGA	1980
Qy	641	LeuLeuAspIlysAlaGlnAspLeuPheGlyAspAspHisAsnIlyAsnGlyPheLysMet	660
Db	1981	CTGCTGGACAAGCCCGAGACCTGTTTGGAGACGACACATAAGAACGGTTCAAAATG	2040
Qy	661	PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheIlyAspAlaThrValArgAla	680
Db	2041	TTGACATCTCTCAACTATCATGGCCAAAGACCTGCTTTTCAAGGATGCCACCGTCCGGCG	2100
Qy	681	ValProValGlyGluIlyThrThrTyrArgGlyTyrPLeuGlyLeuAspTyrValAlaAla	700
Db	2101	GTGCTGTGCGAGAGAAACCACTACCGGGCTGCTGGGCTGGACTACGTGGCGCG	2160
Qy	701	LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGlyAlaPro	720
Db	2161	CTGGAAGGGATGTGCTCTCAGCAGTCTCGGGCGCAGCGCCCCCGCGCGCGCCC	2220
Qy	721	LeuLeuProLeuLeuLeuProAlaLeuAlaAArgLeuLeuProProAlaLeu	738
Db	2221	CTGCTCCCCGTGCTGCTCCCGCCCTCGCGCCCGCTGCTCCCGCCCGCCCTC	2274
RESULT 3			
ABK12569			
ID	ABK12569 standard; cDNA; 2368 BP.		
AC	ABK12569;		
DT	18-JUN-2002 (first entry)		
XX	DNA encoding cell differentiation stimulator associated protein #4.		
XX	Cartilage cell differentiation stimulator; osteopathic;		
XX	Membrane-bound transferrin-like protein; Mtf-BP; concanavalin A; ConA;		
XX	membrane bound type transferrin-like protein; Mtf; cartilage disorder;		
XX	bone metabolism disease; cell differentiation; cell growth;		
XX	extracellular matrix related disease; gene; ss; human.		
OS	Homo sapiens.		
XX	Key		
XX	Location/Qualifiers		
FT	61..2277		
FT	/tag= a		
FT	/product= "Cell differentiation stimulator associated		
FT	protein #4"		
XX	JP2002020311-A.		
PN	XX		
XX	XX		
PD	23-JAN-2002.		
XX	XX		
PF	07-JUL-2000; 2000JP-00206566.		
XX	XX		
PR	07-JUL-2000; 2000JP-00206566.		
XX	XX		
XX	(KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.		
PA	XX		
XX	WPI; 2002-287405/33.		
DR	P-PSDB; AAU78363.		
XX	A cartilage cell differentiation stimulator useful in the diagnosis of		
PT	biophylaxis, cell differentiation, cell growth and construction of		
PT	extracellular matrix related diseases.		
XX	XX		
PS	Disclosure; Page 13-14; 17pp; Japanese.		
XX	XX		

CC The invention describes a cartilage cell differentiation stimulator  
 CC (containing a membrane-bound transferrin-like protein (MTF-BP) and a  
 CC membrane bound type transferrin-like protein (MTf)) and an animal-derived  
 CC concanavalin-like drug. The cartilage differentiation stimulator can be  
 CC used in diagnosis, prevention and treatment of cartilage and bone  
 CC metabolism diseases. They can also be used for diagnosing biophysics,  
 CC cell differentiation, cell growth and construction of extracellular  
 CC matrix related diseases. MTF-BP strongly stimulates differentiation of  
 CC cartilage cells and exhibits similar action mechanism with that of plant  
 CC derived ConA. This sequence represents a cartilage cell differentiation  
 CC stimulator associated polypeptide described in the invention  
 XX

SQ Sequence 2368 BP; 427 A; 766 C; 769 G; 406 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	0	Length:	2368
Score:	3936.00	Matches:	738
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-049-957-4 (1-738) x ABK12569 (1-2368)

QY	1	MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly	20
DB	61	ATGCGGGTCCGAGCGGGGCTGTGGCTGCTCGCCACCGTGTCTCGAGGC	120
QY	21	MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer	40
DB	121	ATGAGGTTGGTGGTGGCGCACCTCGACCCAGAGCAGCACAAAGTGGCGCAACATGAGC	180
QY	41	GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla	60
DB	181	GAGGCTTCGGGAGCGGGCATCCAGCCCTCCTCTGCGTCCGGGCGACCTCCGCC	240
QY	61	AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly	80
DB	241	GACCACTGGGTCAGCTCATCGCGGCCAGGAGGCTGACGCCATCACTCTGGATGAGGA	300
QY	81	AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp	100
DB	301	GCCATCTATAGCGCGGGAAGAGCACCGCCTGAAGCCGGTGGTGGCGCAAGTGTACGAT	360
QY	101	GlnGluValGlyThrSerTyrTrpAlaValAlaValValArgArgSerHisValThr	120
DB	361	CAAGAGTGGTACCTCTATTACGCGGTGGTGTGGTGGAGGAGCTCCCATGTGACC	420
QY	121	IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp	140
DB	421	ATTGACACCTGAAAGGCGTGAAGTCTCTGCCACACGGGCATCAATGCCACAGTGGGCTGG	480
QY	141	AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal	160
DB	481	AACGTCCCGTGGGCTACCTGGTGGAGAGCGCGCCCTCTCGGTGATGGGCTGCGATGA	540
QY	161	LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyValAlaGluThrSer	180
DB	541	CTCAAGCTGTGACGACATATTTTGGGGGAGCTGCTCCCGGGGCGAGAGACCACT	600
QY	181	TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGlyValCysAsp	200
DB	601	TACTCTGAGTCCCTCTGTGCGCTCTCGAGGGGTGACAGCTCTGGGGAGGGGTGTGTAC	660
QY	201	LysSerProLeuGluArgTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly	220
DB	661	AAGAGCCCCCTGGAGAGATACATACGACTACAGCGGGGCTTCGGGTGGCTGGCGAAGG	720
QY	221	AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr	240
DB	721	GCAGGGGAGTGGCTTTTGTGAAGCACAGACCGTACTGGAGAACACCGATGGGAAGCG	780
QY	241	LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly	260

DB	781	CTTCCCTCTCGGGCCAGGCCCTGTGTACAGGACTTCAGCTGTGTGCGGGATGGT	840
QY	261	SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla	280
DB	841	AGCCGGGCGCATGTCAACGAGTGGAGGAGTGCATCTGGCCGGGTGCTGTCTACGCC	900
QY	281	ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln	300
DB	901	GTGGTGGTTCGGGCCGACACATGGGGGCTCACTTCCGGCTGTCTCAACGAGGCCAG	960
QY	301	ArgLeuPheSerHisGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln	320
DB	961	CGTCTGTTCCAGCCAGGCGCAGCTTCAGATGTTTCAGCTCTCAGGCTATGCGCCAG	1020
QY	321	LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr	340
DB	1021	AAGGATCTACTCTTCAAGACTCTTACCTCGGAGCTTGTGCCCATGCCACAGACTAT	1080
QY	341	GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn	360
DB	1081	GAGGCGTGGTGGGCCATGATGCTGCACGCCCATGAAGGCTGTCTGTGACCCCAAC	1140
QY	361	ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly	380
DB	1141	CGGCTGCCCCCTACTCTGGCTGGTGTGTCTCTCCACTCCCGAGATCCAGAAGTGTGA	1200
QY	381	AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla	400
DB	1201	GACATGGCGCTGGCTTCGCCCGGAGCGCCTCAAGCCAGAGATCCAGTGGTGTGAGCC	1260
QY	401	LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu	420
DB	1261	AACTCCCCCAACATGCTGATGGAGCGGATCCAGGCTGACAGCTGACGCTGTGACCTTA	1320
QY	421	SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu	440
DB	1321	AGTGGCGAGGACATTTACACGCGGGGAAGAGTAGTACGGCTCTGTTCCCGCAGCGCGAG	1380
QY	441	HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgAsp	460
DB	1381	CACATATGCCCCGGAAGACAGCAGCAACTCGTACTACGTGGTGGCGTGTGTGAGCGGAC	1440
QY	461	SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe	480
DB	1441	AGCTCCCGGCTTCACTTGGATGAGCTTCGGGCAAGCGCTCTCTGCCACCGCGGTTTC	1500
QY	481	GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg	500
DB	1501	GGCAGCCCTGACAGGCTGGGATGTCCCGTGGGTGCCCTTATTTCAGAGAGGCTTCATCCG	1560
QY	501	ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro	520
DB	1561	CCCAAGGACTGTGAGCTCTCACAGCAGTGAAGGAGTCTTCAATGCCAGCTGCGTGCCC	1620
QY	521	ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln	540
DB	1621	GTGAACAACCCCAAGAACTACCCCTCTCTGCTGTGTGACTGTGGTGGGGAGCGAGCAG	1680
QY	541	GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe	560
DB	1681	GGCCGCAACAAGTGTGTGGGCAACAGCCAGGAGCGGTATTACGGCTACCGCGCGCTTC	1740
QY	561	ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp	580
DB	1741	AGGTCCCTGGTGGAGAAATCCGGGTGACGTGCTTCTCGTCAGGCACACACCGCTCTTAC	1800
QY	581	AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu	600
DB	1801	AACACAAACGGCCCAATTCGAGGCCCTGGGCTGTGAGCTCAGGTGAGGAGACTATGAA	1860
QY	601	LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla	620

```
Db 1861 CTGCTGTGCCCCAACGGGGCCGAGCCGAGGTGTCCAGTTTTCAGCCTGCAACCTGGCA 1920
Qy 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
Db 1921 CAGATACCAACCCAGCCCGGTGATGTCGCGCCCGACACCAACATCTTCCACCGTGTATGGA 1980
Qy 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660
Db 1981 CTGTGGACACAGCCCGAGACCTGTTTGGAGACGACCAACATPAGAACGGGTTCAAATG 2040
Qy 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla 680
Db 2041 TTTCAGCTCTCCAACTATCATGGCCAAAGACCTGCTTTTCAAGCATGCCACCGTCCGGGGC 2100
Qy 681 ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700
Db 2101 GTGCTGTGCGAGAGAAACCACTACCGCGCTGCTGGGCTGGACTACGTGGCGGGC 2160
Qy 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720
Db 2161 CTGGAAGGATGTGCTCAGCAGTGTCTCGGCGGACGCGCCCGCGCGCGCGCGCGCC 2220
Qy 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738
Db 2221 CTGCTCCGCTGTGCTGCGCGCCCTCGCGCGCGCGCTGCTCCCGCGCGCCCTC 2274

RESULT 4
AAN70607
ID AAN70607 standard; DNA; 2362 BP.
AC AAN70607;
XX
DT 25-MAR-2003 (revised)
DT 01-JUL-2002 (revised)
DT 19-FEB-1991 (first entry)
XX
DE Sequence encoding the melanoma-associated p97 antigen.
XX
KW Cancer vaccine; melanoma vaccine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 55..2271
FT FT /*tag= a
FT misc_feature 2362
FT FT /*tag= b
FT FT /note= "1280 BP separates this BP from the 5'-end of the
FT FT SQ (AAN71179)"
XX
PN NL8700285-A.
XX
XX 01-SEP-1987.
XX
XX 06-FEB-1987; 87NL-00000285.
XX
XX 07-FEB-1986; 86US-00827313.
XX 27-JAN-1987; 87US-00007230.
XX (ONCO ) 'ONCOGEN.
XX
XX Brown JP;
XX
XX WPI; 1987-275646/39.
XX P-PSDB; AAP70382.
XX
XX Melanoma-associated antigen peptide(s) - are contained in recombinant
XX PT viruses for prodn. of vaccines against melanoma.
XX
XX Claim 2; Fig 3(A-B); 69pp; Dutch.
XX
XX The sequence encoding the melanoma-associated p97 antigen and recombinant
XX CC viruses contg. it are useful for prodn. of vaccines (claimed) against
XX CC
```

```
CC melanoma. Insertion of the p97 gene into vaccinia virus yields the
CC recombinant virus Vp97a-NY. BSC cells infected with this virus produce
CC large amts. of p97, and the virus has immunogenic activity in vivo.
CC (Updated on 01-JUL-2002 to add missing PI field.) (Updated on 25-MAR-2003
XX to correct PA field.)
SQ Sequence 2362 BP; 427 A; 764 C; 765 G; 406 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2362
Score: 3918.00 Matches: 735
Percent Similarity: 99.59% Conservative: 0
Best Local Similarity: 99.59% Mismatches: 3
Query Match: 99.54% Indels: 0
Db: 1 Gaps: 0

US-10-049-957-4 (1-738) x AAN70607 (1-2362)
Qy 1 MetArgGlyProSerGlyAlaLeuTyrLeuLeuAlaLeuArgThrValLeuGlyGly 20
Db 55 ATGCGGGTCCGAGAGGGGGCTCTGTGGCTGTCTCTGGCTCTGGCCACCGTGTCTCGAGGC 114
Qy 21 MetGluValArgTyrCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40
Db 115 ATGGAGGTGCGGTGTGTGGCCACCTCGGACCCAGAGCAGCAAGTGCAGCAACATGAGC 174
Qy 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60
Db 175 GAGGCTTCCGGGAAGCGGGCATCCAGCCCTCCCTCTCTCGTCCGGGACACCTCCGCC 234
Qy 61 AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80
Db 235 GACCACTCGTCCAGCTCATTCGGGCCCGAGGGCTGACGCATCCTCTCTGGATGAGGA 294
Qy 81 AlalTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100
Db 295 GCCATCTATGAGGGGCAAGAGCAGCGCTGAAGCCGGTGGTGGGCGAAGTGTACGAT 354
Qy 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValAlaValArgSerSerHisValThr 120
Db 355 CAAGAGGTGCGGTACCTCTCTATTACGCCGTGTGTGTGAGAGGAGTCCCATGTGACC 414
Qy 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140
Db 415 ATTGACACCTTGAAAGGCGTGAAGTCTCTGCCACCGGCATCATCGCACAGTGGGCTGG 474
Qy 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160
Db 475 AACGTGCCCGTGGGCTACCTGTGTGAGAGAGCGCGCCCTCTCGGTGATGGGCTGCATGTA 534
Qy 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180
Db 535 CTCGAAGCTGTACGAGCTATTTTGGGGGAGCTGCGTCCCGGGGCGAGGAGACCAAGT 594
Qy 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200
Db 595 TACTCTGAGTCCCTCTGTCTGCTCTGTGAGGGGTGACAGCTCTGGGGAAGGGGTGTGTAC 654
Qy 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly 220
Db 655 AAGAGCCCTCGGAGAGATACCTAGACTACAGCGGGGCTTCCGGTGTCTGGCGAAGGG 714
Qy 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240
Db 715 GCAGGGGACGTGGGCTTTTGTGAAGCACAGACGGTACTGGAGAACACCGATGGGAAGAGC 774
Qy 241 LeuProSerTyrGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260
Db 775 CTTCCCTCTGGGGCCAGGCCCTGCTGTCTCAGAGACTTCGAGCTGTCTGCGGGATGGT 834
Qy 261 SerArgAlaAspValThrGluTyrArgGlnCysHisLeuAlaArgValProAlaHisAla 280
Db 835 AGCCGGGCGGATGTACCGAGTGGAGGAGTGCCATCTGTGGCCCGGGTGCCTGCTCACGCC 894
```



QY	281	valValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln	300	QY	641	LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet	660
DB	895	GTGGTGGTCCGGCCGACACAGATGGGGCTCATCTTCGGGTGCTCAACGAAGCCAG	954	DB	1975	CTGCTGGACAAGGCCCAGGACCTGTTTGGAGACGACCAATAAGAACGGTTCAAATG	2034
QY	301	ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln	320	QY	661	PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla	680
DB	955	CGTCTGTTACGCCACAGGGCAGCAGCTTCCAGATGTTTCCAGCTCTCAGGCTATGGCCAG	1014	DB	2035	TTTGACTCTCTCAACTATCATGCGCAAGACCTGCTTTTCAAGGATGCCACCGTCCGGCG	2094
QY	321	LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr	340	QY	681	ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla	700
DB	1015	AAGGATCTACTTTCAGAGCTCTACTCGAGCTTGTGCCATCGCCACACAGACTAT	1074	DB	2095	GTGCTGTTCGGAGAGAAACACCTACCGCGGCTGGTGGGCTGGACTACGTGGCGCG	2154
QY	341	GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn	360	QY	701	LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro	720
DB	1075	GAGCGTGGTGGCCCATGAGTACCTGACCGCATGAAGGGTCTGCTCTGCCACCAAC	1134	DB	2155	CTGGAAGGATGCTGCTCAGCAGTGTCTCGGCGCAGCGGCCCGCGCGCGCGGCC	2214
QY	361	ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly	380	QY	721	LeuLeuProLeuLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu	738
DB	1135	CGGCTGCCCCCTTACCTGGCGTGGTGTGTCTCTCCACTCCCGAGATCCAGAAGTGTGA	1194	DB	2215	TTGCTCCCGCTCTGCTGCCCGCCCTCGCGCGCGCTGCTCCCGCGCGCCCTC	2268
QY	381	AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla	400	RESULT 5			
DB	1195	GACATGGCGTGGCTTTCGCCGGCAGCGCTCCAGCCAGAGATCCAGTGGTGTGCGCC	1254	ID	AAF63460	standard; cDNA; 2154 BP.	
QY	401	LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu	420	AC	AAF63460;		
DB	1255	AGTCCCCCAACACACTGCATGGAGCGGATCCAGGCTGAGCAGTTCGACCTGTGAC	1314	DT	XX		
QY	421	SerGlyGluAspIleTyrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu	440	XX	11-MAY-2001 (first entry)		
DB	1315	AGTGGCGAGACATTTACACGGCGGGAAGTACGGCTGTTTCCCGCAGCGCGCGAG	1374	XX	Human MTF coding sequence.		
QY	441	HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgAsp	460	XX	Human; MTF; arthritic disorder; melanotransferrin; rheumatoid arthritis;		
DB	1375	CACATGTCGCCGGAACAGACAGCAACTCGTACTAGTGGTGGCGGTGTGAGACGGAC	1434	KW	osteoarthritis; joint trauma; ss.		
QY	461	SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe	480	XX	Homo sapiens.		
DB	1435	AGCTCCACCGCTTACCTTGGATGAGCTTCGGGCGAAGCGCTCTGCCACCGCGGTTTC	1494	XX	WO200111368-A1.		
QY	481	GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheLeuArg	500	XX	15-FEB-2001.		
DB	1495	GGCAGCCCTCGAGCTGGGATGTCCTCGTGGTGGTGGTCTTATTTCAGAGAGCTTCATCCG	1554	XX	02-AUG-2000; 2000WO-JP005171.		
QY	501	ProLysAspCysAspValLeuThrAlaValSerGluPheAsnAlaSerCysValPro	520	XX	05-AUG-1999; 99JP-00222568.		
DB	1555	CCCAAGGACTGTGACGTCTCACAGCAGTGAGCGAGTCTTCAATGCCAGCTCGGTGCC	1614	XX	(ORIV ) ORIENTAL YEAST CO LTD.		
QY	521	ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln	540	XX	Kato Y, Matsukawa H, Yoshiwara Y, Oka O, Fujita T;		
DB	1615	GTGAACAACCCCAAGAACTACCCCTCTCGTGTGTGCTGTGCTGGTGGGACGAGCAG	1674	XX	WPI; 2001-191574/19.		
QY	541	GlyArgAsnLysCysValGlyAsnSerGlnArgTyrTyrGlyTyrArgGlyAlaPhe	560	DR	P-PSDB; AAB73169.		
DB	1675	GGCCGCAACAAGTGTGGGCAACAGCCAGGAGCGGTATTACGGCTACCGCGCGCTTC	1734	PT	Assay of melanotransferrin by sandwich immunoassay in blood or synovial		
QY	561	ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp	580	PT	fluid for diagnosis of arthritic disorders.		
DB	1735	AGTGTCTGTGGAGAAATCGGTGTGCTTCCGTGAGGACACCAACCGCTCTTTC	1794	XX	Disclosure; Page 30-38; 40pp; Japanese.		
QY	581	AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu	600	XX	The present invention relates to a method for the diagnosis of arthritic		
DB	1795	AACACAAAAGCCACAAATTCGAGCCCTGGGCTGTGCTGAGCTCAGGTACAGAGACTATGA	1854	CC	disorders, by assay of melanotransferrin in a biological sample, and		
QY	601	LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla	620	CC	deducing the presence of an arthritic disorder from the presence of a		
DB	1855	CTGCTGTGCCCAACAGCGGCCCGAGCGAGGTGTCCAGTTTGCAGCTTGCACCTGGCA	1914	CC	concentration of melanotransferrin significantly higher than in non-		
QY	621	GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly	640	CC	arthritic subjects. The method can be used for the diagnosis of arthritic		
DB	1915	CAGATACCAACCCACCGCTGATGGTCCGCGCCGACACCAACATCTTCCCGTGTATGA	1974	CC	disorders including rheumatoid arthritis, osteoarthritis and joint		
				CC	trauma. The present sequence is the coding sequence for human MTF, which		
				CC	was used in the present invention		
				XX	Sequence 2154 BP; 402 A; 677 C; 695 G; 380 T; 0 U; 0 Other;		
				SQ			
				Alignment Scores:			
				Pred. No.:	0	Length:	2154
				Score:	3834.00	Matches:	718
				Percent Similarity:	100.00%	Conservative:	0
				Best Local Similarity:	100.00%	Mismatches:	0
				Query Match:	97.41%	Indels:	0
				DB:	4	Gaps:	0

US-10-049-957-4 (1-738) x AA#63460 (1-2154)

QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
DB 1 ATGGAGGTGGTGGTGGCCACCTCGGACCCAGAGCAGCAAGTGGCGCAACATGAGC 60  
QY 41 GluAlaPheArgGluAlaGlyLeuGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
DB 61 GAGGCGCTCCGGGAAGCGGCGATCCAGCCCTCCCTCTGGTCCGGGCGACCTCCGCC 120  
QY 61 AspHisCysValGlnLeuLeuAlaGlnGlnAlaAspAlaIleThrLeuAspGlyGly 80  
DB 121 GACCACTGGTCCAGCTCATCCGGCCAGAGGCTGACGCCATCACTCTGGATGGAGA 180  
QY 81 AlaIleThrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTrpAsp 100  
DB 181 GCCATCTATGAGCGGGAAGAGCAGCGCCCTGAAGCGCGTGGTGGCGAAGTGTACGAT 240  
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValArgArgSerSerHisValThr 120  
DB 241 CAAGAGGTCCGTACCTCTATTACCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300  
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp 140  
DB 301 ATTGACACCTGAAGGCGTGAAGTCTGCGCACACGGGCATCAATCGCACAGTGGCTGG 360  
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
DB 361 AACGTGCCCGTGGGTACCTGGTGGAGAGCGCGCCCTCTCGGTGATGGGCTGCGATGTA 420  
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyClnThrSer 180  
DB 421 CTCGAAGCTGTCCAGCGACTATTTGGGGCAGCTCGTCCCGGGGCGAGGAGCCAGT 480  
QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200  
DB 481 TACTCTGAGTCCCTCTGTGCGCTTCGAGGGGTGACAGCTCTGGGGAGGGGTGTGTGAC 540  
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly 220  
DB 541 AAGAGCCCCCTGGAGAGATACACGACTACAGCGGGGCTTCGGGTGCTGGCGGAAGGG 600  
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
DB 601 GCAGGGGAGTGGCTTTTGTGAAGCACAGCAGCGGTACTGGAAACACGGATGGGAAGACG 660  
QY 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260  
DB 661 CTTCCCTCTGGGGCCAGGCCCTGTGTACAGGACTTCGAGCTGTGTCGGGATGGT 720  
QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
DB 721 AGCGGGGCGGATGTCAACGAGTGGAGCAGTGCCTATCTGGCCCGGGTGCCTGCTCACGCC 780  
QY 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300  
DB 781 GTGGTGGTCCGGGCGACACAGATGGGGCCCTCATCTTCGGGCTGTCTCAACGAAGGCCAG 840  
QY 301 ArgLeuPheSerHisGluGlySerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
DB 841 CGTCTGTTCCAGCCAGGCGGAGCGTTCAGATGTTTCAGCTCTGAGGCGCTATGGCCAG 900  
QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
DB 901 AAGGATCTACTCTTCAAGACTCTACTCGGAGCTTGTGCCCATCGCCACACAGACCTAT 960  
QY 341 GluAlaTrpLeuGlyHisGlnTyrLeuHisAlaMetLysGlyLeuLeuCysAspProben 360  
DB 961 GAGGCGTGGCTGGGCCATGAGTACCTGACCGCCATGAAGGGTCTGCTCTGTGACCCCAAC 1020  
QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380

DB 1021 CGGTGCCCCCTACTGCGCTGTGTGTCTCTCACTCCCGAGATCCAGAAAGTGTGGA 1080  
QY 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
DB 1081 GACATGGCGGTGGCTTTCCCGCGCAGCGCTCAAGCCAGAGATCCAGTGGTGTACGCC 1140  
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
DB 1141 AAGTCCCCCAACTGCTGATGAGCGGATCCAGGCTGAGCAGGTCGACGCTGTGACCTTA 1200  
QY 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
DB 1201 AGTGGCAGGACATTTTACACGCGCGGAAGAAGTACGCGCTGGTTCCTCCGACGCGCGAG 1260  
QY 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp 460  
DB 1261 CACTATGCCCCGGAAGACAGCAGCAACTCGTACTACGTGGTGGCGTGGTGGTGGTGG 1320  
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
DB 1321 AGCTCCACGCTTCACTTGGATGAGCTTCGGGGCAGCGCTCTGCGCACCGCGGTTC 1380  
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
DB 1381 GGCAGCCCTGCAGGCTGGGATGTCCCGTGGGTGCCCTTATTCAGAGAGGCTTCATCCGG 1440  
QY 501 ProllysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
DB 1441 CCCAAGGACTGTGAGCTCTCACAGCAGTGGAGCGATTCTTCAATGCCAGCTGGTGGCC 1500  
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
DB 1501 GTGAAACACCCAGAGACTACCTCTCTCGTGTGTGCACTGTGCGTGGGGGACGAGGAG 1560  
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
DB 1561 GGCGCAACAAGTGTGTGGGCAACAGCAGCAGAGCGGTATTACGGCTACCGCGGCGCTTC 1620  
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrValPheAsp 580  
DB 1621 AGGTGCTGTGGAGAAATGCGGGTGAGCTGCCCTTGTGTCAGGCACACAACTCTTTGAC 1680  
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600  
DB 1681 AACCAACAGCCACATTCGAGCCCTGGGCTGCTGAGCTCAGGTGAGAGGACTATGAA 1740  
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
DB 1741 CTGCTGTGCCCAACGCGGCGCGAGCGGAGGTGTCCAGTTCAGAGCTGCAACCTGGCA 1800  
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
DB 1801 CAGATACACCCCGCGCGTGTGTCGGGCGCGACACCAACATCTTCACCGTGTATGGA 1860  
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660  
DB 1861 CTGTGGACAGGCCAGACCTGTTGGAGACGACCAACATAGAACGGGTTCAAATG 1920  
QY 661 PheAspSerSerAsnTyrHisGlyClnAspLeuPheLysAspAlaThrValArgAla 680  
DB 1921 TTCGACTCTCCAACTATCATGGCCAGACCTGCTTTTCAAGGATGCCACCGCTCGCGGCG 1980  
QY 681 ValProValGlyGluLysThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaAla 700  
DB 1981 GTGCTGTGCGAGAGAAACCACTTACCGCGGCTGGCTGGGCTGGACTAGTGGCGGCG 2040  
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720  
DB 2041 CTGAGAGGATGTGCTCTCAGCAGTGTCTGGGCGCAGCGGCGCGCGCGCGCGCGCGCC 2100  
QY 721 LeuLeuProLeuLeuProAlaLeuAlaArgLeuLeuProProAlaLeu 738  
DB 2101 CTGCTCCGCTGCTGCTGCCCGCGCTGCGCGCGCGCTGCTCCGCGCGCGCGCTC 2154





QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValAlaValArgSerSerHisValThr 120  
DB 360 CAAGAGGTGGGACCTCTACTACGTGTGGCGGTGTCAAGAGGAGCTCCAAACGTGACC 419  
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140  
DB 420 ATCAACACCTGAGAGCGTGAAGTCTGTGCCACACGGGCAACACGCGGTGGGTGG 479  
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
DB 480 AACGTGCTGTGGGTACCTGTGGTGGACAGCGCGCGCTCTCACTGATGGGCTGTGACGTG 539  
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180  
DB 540 CTCAAGACGCTACGAGGACTTTCGGGGGCGAGTGGTCCCTGGGGGAGAGACAGACA 599  
QY 181 TyrSerGluSerLeuLysCysArgLeuLysCysArgGlyValSerSerGlyGlyValCysAsp 200  
DB 600 TACTCGAGTCCCTCTGTGCTCTGTCCGGGGGCGACACCTCCGGGGGAGGGGTGTGTGAC 659  
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGlyGly 220  
DB 660 AAGAGCCCTGTGGAGCGGTACTACGACTACAGCGGGGCTTCGGGTGCTTGGCAGAGGC 719  
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyTyrThr 240  
DB 720 GCAGGGGAGTGGCTTTGTGAAGCACAGCACGGTGTGTGAGAACACGGATGGGAGAAC 779  
QY 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLysCysArgAspGly 260  
DB 780 CTGCTCTCTGGGGGCGACATGCTGATGTACGGGACTTTGAGTGTGTGTGGCGGACGGC 839  
QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
DB 840 AGCGGGGCGAGCTCACCGAGTGGAGACACTGCCACTTGGCCGGGTGGCCGCCACGCC 899  
QY 281 ValValAlaArgAlaAspThrAspGlyGlyLeuLysPheArgLeuLeuAsnGluGlyGln 300  
DB 900 GTGGTGTGTGGGGCGGACACCGACCGAGCGCTCATCTCCGGTCTCTCAATGAGGGCCAG 959  
QY 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
DB 960 CGGCTGTTCAGCCACGAGGGGACAGCTTCCAGATGTTCCAGTCTCGAGGCGCTACGGCCAG 1019  
QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
DB 1020 AAGAACCCTGCTGTTCAGAGCTCCACGCTGGAGCTGGTGCCCATCGCCACACAGACTAC 1079  
QY 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuLysAspProAsn 360  
DB 1080 GAGGCTGTGGGGCGGCGGAGTACCTGTCCACCGCATGAAGGTCTGTCTGTGACCCCAAC 1139  
QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
DB 1140 CGGCTGCCCGCCACFACCTGGCTGGTGGTGTGTCCACCCCGAGATCCAGAGGTGGGA 1199  
QY 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
DB 1200 GACATGGCGGTGGCTTTCAGCGCGGAGAGGCTCAAGCCGGAGATCCAGTGTCTCGGGC 1259  
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
DB 1260 GAGTCCCGCCAGCACTGCATGAGAGCAGATCCAGGCTGGGCGACATCGATGTGTGACCCCTG 1319  
QY 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
DB 1320 AACGGGGAGGACATTCACACAGCGGGGAGAGCTTATGGCTGATCCCGGCTGGCGGGAG 1379  
QY 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValAlaValArgArgAsp 460  
DB 1380 CTGTATGCCCGGACACAGGAGTAACCTGTACTTCTGTGGTGGCGGTGTGTGAAGCGAGAC 1439

QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
DB 1440 AGCGCTACGCTTACCGTCGAGCGAGCTGCGCGGAGCGCTCTCTGCCACCGCGCTTC 1499  
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
DB 1500 GGCAGCGCGCGCGCTGGGACGTCCCGGTGGCGCCCTCATCCACTGGGGCTACATCCGG 1559  
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
DB 1560 CCCAGGAACCTGCAGCGTCTCACAGCGTGGGTTCAGTCTTCTCAACGCGAGCTGTGCGCG 1619  
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
DB 1620 GTGAACAACCCCAAGATACCCCTCTCGCTGTGCGCACTCTGCGTGGGTGACGAGCAG 1679  
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
DB 1680 GGCCGCAACAGTGTCACTGGCAACAGCCAGGAGCGGTACTATGGCGACAGTGGCGCTTC 1739  
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
DB 1740 AGTGCCTGTGTGAGGGTGCAGGGGACGTGGCTTCGTCAAGCACACAGCACTCTTGAC 1799  
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600  
DB 1800 AACCAAAATGGCCACAATCCCGAGCGGTGGCTGCCATCTGAGGAGCCAGGACTACGAG 1859  
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
DB 1860 CTGCTGTGCCCCAACCGCGCGGAGCTGAGGCGCCACAGTTTGGCGCCCTGCAACCTGGCC 1919  
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
DB 1920 CAGATTCCGTCCACCGCGTCACTGTGGCGCCCGACCAACATCTTCAACCGTTACGGA 1979  
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660  
DB 1980 CTGTGGACCAAGCGCCAGGACCTGTGTGGAGACGACCAACAAGAACGGGTCAAGATG 2039  
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680  
DB 2040 TTCGACTCTCTCAGCTTACCACCGCGGAGACCTGTCTTCAAGAGCGCCAGTGGCGGCT 2099  
QY 681 ValProValGlyLysLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700  
DB 2100 GTGCTGTGGCGGAGAGGACCACTTACAGGACTGGCTGGGCGCGGACTACGTGGCGCT 2159  
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720  
DB 2160 CTGGAAGGGATGCACTCAGCGTGTCTCAGGGGCGAGCGCTCGGGCGCCCGCGGCGCTCG 2219  
QY 721 LeuLeuProLeuLeuProAlaAlaArgLeuLeuProAlaLeu 738  
DB 2220 CTGCTGGCGCTGCTG-----CCCTGTGGTGGCGGCTCTCTGTCTTCTGCTC 2267  
RESULT 8  
AAD14466  
ID AAD14466 standard; cDNA; 4068 BP.  
XX AAD14466;  
XX  
XX 01-NOV-2001 (first entry)  
XX  
XX Mouse p97 (mp97) protein cDNA.  
XX  
KW Mouse; mp97 protein; scioglycoprotein; neuroprotective; antibacterial;  
KW analgesic; nootropic; cytostatic; neuroleptic; virucide; anticonvulsant;  
KW deficiency disease; Wernicke's disease; neurodegenerative disease; pain;  
KW nutritional polyneuropathy; neurological disorder; cancer; gene therapy;  
KW Huntington's disease; Alzheimer's disease; Parkinson's disease; epilepsy;  
KW demyelinating disease; multiple sclerosis; amyotrophic lateral sclerosis;  
KW psychosis; therapeutic; ss.



Db 784 CTGCTCTCCCTGGGGCAAGTCCCTGATGTGTGACAGGACTTCCAGCTACTATGACGGGATGCG 843  
QY 261 SerArgAlaAspValThrGluTyrArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
Db 844 AGCCGAGCCGACATCACTGATGGAGAGCTGTGCACCTGGCCAAAGTGTCTGCTCATGCT 903  
QY 281 ValValValArgAlaAspThrAspGlyGlyLeuLeuPheArgLeuLeuLeuLeuGluGlyGln 300  
Db 904 GTGGTGTGAGGGGTGACATGATGATGCGGTCTCATATTCACATGCTCAACGAAGCCAG 963  
QY 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
Db 964 CTTTCTGTTGAGCCAYGAAGACAGCAGCTTCCAGATGTTCCAGTCCAAAGCCTACAGCCAG 1023  
QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
Db 1024 AGAACTTGTCTGTTCAAGACTCCACCTTGGAGCTTGTGCCATTGCCCACACAGAATAT 1083  
QY 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuLeuLysAspProAsn 360  
Db 1084 GAGGCTGTGGTGGCCAGGAATACCTGCAGGCCATCAAGGGGCTCTCTGTGATCCCAAC 1143  
QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
Db 1144 CGGCTGCCCTACCTGCTGCTGGTGTGCTGTGCTGACGCGCCGAGATCCAGAGTGTGA 1203  
QY 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
Db 1204 GATATGGCTGTGGCTTTCAGCCGCCAGAAATCTCAAGCCGGAATTCAGTGTGTGTCGGCC 1263  
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
Db 1264 GAGTCCCTGAGCACTGCATGGAGCAGATCCAGGCTGGGCACACTGACGCTGTGACTGTG 1323  
QY 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
Db 1324 AGGGGCGAGGACATTTACAGGGCAGGAAAGGTGTACGGCTGTGTTCCGGCGCGGGGAG 1383  
QY 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp 460  
Db 1384 CTGTATGCTGAGGAGACAGGAGCAATTCCTACTTTGTGTGCTGTGGCAAGAGGAC 1443  
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
Db 1444 AGCTCTACTCTTCACTCCCTGGAGCAGCTTGGCGCAAGGTTCTTCGCCACCTACTTG 1503  
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
Db 1504 GGCAGCCAGCGGGCTGGAGGTGGCCATCGGCTCCCTCATCCAGCGGGGCTTCATCCGG 1563  
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPheAsnAlaSerCysValPro 520  
Db 1564 CCCAAGACTGTATGTCTTCACGCGGTGAGCCAGTTCTTAATGCCAGCTGCGTGCCT 1623  
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
Db 1624 GTCAACAACCTTAAGAACTACCTTCCGCACTATGTGCGCTCTCGTGGGAGACGAGAAG 1683  
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
Db 1684 GGCCGCAACAAATGTGTGGGAGACGCCAGGAGATACCTACGGCTTACAGCGGGGCTTTC 1743  
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
Db 1744 AGGTGCTGTGTGAGCATGAGGGAGCGTGTCTTCGTAAGCACACAGCTGCTTTGAG 1803  
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600  
Db 1804 AACACAAATGGTCACAACTCTGAGCTTGGGCTTCTCACTCAGGTGGCAAGACTATGA 1863  
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
Db 1864 CTACTGTGCCCCAATGGGCGCAGGGCTGAGGTAGACCAAGTTCACAGCTTCAACCTGGCA 1923

QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
Db 1924 CAATGCCATCCACCGCTGTATGTCGTCCAGACCAACATCTTCACTGTGTATGA 1983  
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660  
Db 1984 CTTCTGGACAAGCCCGAGACCTGTTTGGAGACGACCAATAACAAGAACGGTTTCCAATG 2043  
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680  
Db 2044 TTTGACTCTCTCAATATACAGCCCAAGACCTGCTTTTCAAAGATGCTACAGTCCGAGCG 2103  
QY 681 ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700  
Db 2104 GTGCAGTCCGGGAGAAAACCACTACCTGGACTGGCTGGGTCTGACTATGTGGTTGCG 2163  
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720  
Db 2164 CTGGAGGGGATGTTGCTCAGCAGTGTCTCCGTCGAGGGCGCGTCCGAGCGAGTCCCC 2223  
QY 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738  
Db 2224 CTGTGGCCCTGCTCTGCTGACCTGGCTGAGGCTCTTCTCTCGCGTTCTC 2277

## RESULT 9

AAF62198

ID AAF62198 standard; DNA; 4158 BP.

XX AAF62198;

XX 21-MAY-2001 (first entry)

XX DNA encoding membrane bound transferrin like protein.

XX Chondrogenesis promoter; membrane-bound transferrin-like protein; Mtf;

KW Chondrogenesis regulator; Mtf activator; bone metabolism; mouse;

KW Chondral differentiation inhibitor; bone disease; ds.

XX Mus sp.

XX WO200113951-A1.

XX 01-MAR-2001.

XX 21-AUG-2000; 2000WO-JP005590.

XX 19-AUG-1999; 99JP-00232966.

XX (CHUS ) CHUGAI SEIYAKU KK.

XX Kato Y, Fujimoto K;

XX WPI; 2001-218409/22.

XX P-PSDB; AAB62882.

XX Chondrogenesis promoters containing membrane-bound transferrin-like

XX protein, useful in diagnosis, prevention and treatment of diseases due to

XX abnormal chondral metabolism and bone metabolism.

XX Disclosure; Page 43-46; 57pp; Japanese.

XX This invention relates to chondrogenesis promoters containing a membrane-

XX bound transferrin-like protein (MTF). Chondrogenesis promoters,

XX chondrogenesis regulators, Mtf activators, Mtf antagonist-containing

XX chondral differentiation inhibitors are useful in diagnosis, prevention

XX and treatment of diseases due to abnormal chondral metabolism and bone

XX metabolism e.g. bone diseases. The present sequence represents murine DNA

XX encoding Mtf

SQ Sequence 4158 BP; 960 A; 1159 C; 1177 G; 862 T; 0 U; 0 Other;

Alignment Scores:



Pred. No.: 2,13e-285 Length: 4158  
Score: 3300.00 Matches: 618  
Percent Similarity: 89.84% Conservative: 45  
Best Local Similarity: 83.74% Mismatches: 75  
Query Match: 83.84% Indels: 0  
DB: 4 Gaps: 0

US-10-049-957-4 (1-738) x AAF62198 (1-4158)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20  
DB 118 ATGAGGCTCCTGAGCGTACTTTTGGCTACTCTCTGCTCGCCTCGCAGCTGCTGCTG 177  
QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
DB 178 ATGGAGGTCCAGTGTGTACCATCTCAGACGAGAGCAGCAGAGTGCAGAGGATGAGC 237  
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
DB 238 GAGGCTTCCAGGGAGTGGCAATTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 297  
QY 61 AspHisCysValGlnLeuIleAlaLaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80  
DB 298 GACCACTGTCTCAGCTCATCAAGGAACAAAGAGCAGATGTCATCACCTGCTGAGGG 357  
QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100  
DB 358 GCATCTCTATGAGCGAGGAGAGAGCAGCGCCCTGAAGCCAGTGTGGGGAAGTCTATG 417  
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValAlaValArgSerHisValThr 120  
DB 418 CAAGCATTTGGGACTTCTATTATGCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 477  
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp 140  
DB 478 ATCAACACCCTCAAGGGCGTCAAGTCTGCTCCACACAGGCAATTAACGGAGTGTGG 537  
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetCysAspVal 160  
DB 538 AACGTGCTGTGGTGTACCTCTAGAGAGCGCCATCTGTCTGATGTGGCTGTGTGTGT 597  
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyValaGlyGluThrSer 180  
DB 598 CTCGAAGCGTGTGGTATTATTTTGGAGCAGCTGTCTCTCTCTCTCTCTCTCTCTCT 657  
QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200  
DB 658 CATTCGAGTCCCTCTGTGCTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 717  
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGly 220  
DB 718 AAGAGTCCCTAGAGAGATCTACCACTACGTGAGAGCTTCCGCTGCTGCTGCTGCTG 777  
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
DB 778 GCGGTGAGTGGCTCTCGTGAAGCAGCAGCAGTGTGTGGAATACTGTATGGAACACC 837  
QY 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260  
DB 838 CTGCTCTCTGCGGCAAGTCCCTGATGTCTGAGAGGACTTCCAGCTACTATGAGGGATG 897  
QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
DB 898 AGCCGAGCGCATCATCTAGTGGAGAGCTTCCACCTGAGGAGCTTCCGCTGCTGCTGCT 957  
QY 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300  
DB 958 GTGGTGTGTCAGGGGTGACATGATGGGTCTCATATTCCTCACTGCTCAACGAGGCCAG 1017  
QY 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
DB 1018 CTTCTGTTTCAGCCAGACAGCAGCTTCCAGATGTTTCAGCTCCAAAGCCTTACAGCCAG 1077

QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
DB 1078 AAGAACTTGTCTGTTCAAGACTCCACCTTGAGGCTTGTGCCATTGTCCACACAGAACTAC 1137  
QY 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360  
DB 1138 GAGGCTGTGCTGGGCAGGAATACCTGAGGCCATGAAGGGGCTCTCTGTGTATCCCAAC 1197  
QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
DB 1198 CGGCTGCCCACTACTCGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1257  
QY 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
DB 1258 GATATGGCTGTGGCTTTCAGCCGCCAGAAATCTCAAGCCGGAATTCAGTGTGTGTGTGT 1317  
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
DB 1318 GAGTCCCTGAGCACTGATGGAGCAGATCCAGGCTGGGCACACTGACGCTGTGACTGTG 1377  
QY 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
DB 1378 AGGGCGAGGACATTTACAGGCGAGAAAGGTGTACGGCTGTGTGTGTGTGTGTGTGTGTGT 1437  
QY 441 HisTyrAlaProGluAspSerSerAspSerTyrTyrValValAlaValAlaValArgAsp 460  
DB 1438 CTGTATGTGTGAGGAGGACAGGAGCAATTCCTACTTTGTGTGTGTGTGTGTGTGTGTGT 1497  
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
DB 1498 AGCTCTACTCTCTCACCCTGAGCAGCTTTCGGGCAAGCGTCTCTGCCACCCCTACTTTG 1557  
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
DB 1558 GGCAGCCAGCGGGTGGGAGTGGCTCCCTCATCGCTCATCGCGGGCTTCATCCGG 1617  
QY 501 ProlLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
DB 1618 CCCAAGGACTGTGTGTCTCTCACAGCGGTGAGCCAGTCTTCAATGCCAGCTGCTGTGCT 1677  
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
DB 1678 GTCAACACCCTAGAACTACCTTCCGCACTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1737  
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
DB 1738 GCGCCCAACAAATGTGTGGGAGCAGCCAGGAGAGATACCTACGGCTACAGCGGGGCTTC 1797  
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
DB 1798 AGGTGCTTGTGGAGCATGCGGGAGCTGTGCTTTCGTCAGACACCACTGTCTTTGAG 1857  
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600  
DB 1858 AACCAATGTGTCAATCTGAGCTTGGGCTTCTCACTCAGGTGCGAGCACTATGAA 1917  
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
DB 1918 CTACTGTGCCCAATGTGGGCGCAGGCTGAGGTAGACCACTTCCAAAGCTTGCACACTGGCA 1977  
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
DB 1978 CAATGTCCATCCACGCTGTATGTCCGTCCAGACCAACATCTTCACTGTGTGTGTGTGT 2037  
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660  
DB 2038 CTTCTGCAAGGCCAGGCTGTTTGGAGAGCAGCCATAACAGAACCGTTTCCAAATG 2097  
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla 680  
DB 2098 TTTGACTCTCCAAATATACAGGCAAGACCTGTCTTTTCAAGAGATGCTACAGTCCGAGCG 2157  
QY 681 ValProValGlyGluLysThrThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaAla 700



```

Db      2158 GTCCAGTCCGGAGAAAACCATCTGGACTGGCTGGCTCCTGACTATGGTGGC 2217
QY      701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720
Db      2218 CTGAGGGGATGTTGTTCTAGCAGTCTCGGTGCGAGGGCGCGGTGCGAGGATCCCC 2277
QY      721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738
Db      2278 CTGCTGGCCCTGCTCCTGCTGACCTGCTGAGGCTCCTTCTCGCGTCTC 2331

```

## RESULT 10

AAQ84700

ID AAQ84700 standard; cDNA; 2347 BP.

XX AC AAQ84700;

DT 25-MAR-2003 (revised)

DT 09-NOV-1995 (first entry)

XX DT

XX DT

XX DE

XX Draculin cDNA.

XX Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

KW factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

KW acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

KW unstable angina; transient ischemic attack; peripheral vascular; DFP;

KW bypass occlusion; disseminated intravascular coagulation; ds.

XX OS

XX Desmodus rotundus.

XX XX

XX Key

XX Location/Qualifiers

XX CDS

XX 1..1961

XX /\*tag= a

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

Query Match: 33.14%

DB: 2

US-10-049-957-4 (1-738) x AAQ84700 (1-2347)

QY 7 AlaLeuTrpLeuLeuLeuAlaLeuArgThrValLeuGlyGlyMetGlu-----ValArg 24

Db 19 GCACGTCTGTCCTCTCTGCGCCCTGCGCGAGTCTTGGCTGCCGCGAGAGAGGTGTTCGA 78

QY 25 TrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSerGluAlaPheArg 44

Db 79 TGTGTGACCATATCAAAACACAGAGCAGCAAAATGCTTAAACTGCAACAGAAATCTAAA 138

QY 45 GluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAlaAspHisCysVal 64

Db 139 CGAGTGGCGT---GGCCCTCTCTCTCTGATAGCAGAAAGTCTTACCTACCTGGAATGTATC 195

QY 65 GlnLeuIleAlaAlaGlnGlnAlaAspAlaIleThrLeuAspGlyGlyAlaIleTrpGlu 84

Db 196 CAGGCCATCGCGCGCAAAAGGCGCATGCCATGAGCCTTTGATGCAAGTCTGGTGTACGAG 255

QY 85 AlaGlyLysGlu---HisGlyLeuLysProValValGlyGluValTrpAspGlnGluVal 103

Db 256 GCAGACAGCAGCCCGTACAGATTGGCGCTGTGGCAGCAGAGAGTCTACCGGACCGAGGG 315

QY 104 Gly-----ThrSerTyrTrpAlaValAlaValArgSerSerHisValThrIle 121

Db 316 GCACCGCGGACGCACTATTACGCTGTGGCCCTGGTGAAGAGCAGCAACTTGCACCTG 375

QY 122 AspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrpAsn 141

Db 376 AGCCAGCTGCAAGGCGTGAAGTCTGCCACACTGCGCTCAACAGGTCGCCCGGGTGA 435

QY 142 ValProValGly-----TyrLeuValGluSerGlyArgLeuSerValMetGly 157

Db 436 ATCCCTGTGGGACCGCTCGTCCGTACCTGGGCTGGGAGGCGCCACCTCGA----- 486

QY 158 CysAspValLeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGly 177

Db 487 ---CCCTCCAGGAAGTGTGGCAACTTCTTCTCCGCTAGCTGTGTCCCTGTGCAGAT 543

QY 178 GluThrSerTyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGly 197

Db 544 GGCACACAGTAC---CCCAACCTGTGTGCTGTGTGGGGG-----ACAGGGGCGAGAT 594

QY 198 ValCysAspLysSerProLeuGluArgTyrTrpAspTyrSerGlyAlaPheArgCysLeu 217

Db 595 AAATGTGCTGCTCTCCAGGAACCGTACTTTGGCTACTCCGGTGCCTTCAAGTGTCTG 654

QY 218 AlaGluGlyAlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAsp 237

Db 655 AAAGATGGGCTGGAGACGTGGCTTTTGTCAAGGACAGTACCGTGTTCGAGAAC----- 708

QY 238 GlyLysThrLeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCys 257

Db 709 -----TGCCAAACAAAGGCCGAG-----AGAGACCAAGTATGAGTGTCTCTGC 750

QY 258 ArgAspGlySerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValPro 277

Db 751 CCAGAACACACCCGAAGCCGGTGGATGAGTTTGACAGTGCACCTGCGCGGGTCCCT 810

QY 278 AlaHisAlaValValAlaArgAlaAspThrAspGlyGly-----LeuIlePheArg 294

Db 811 TCTCATGAGTTGTGGCCCGA-----AGCGTGGGTGGCAAGGAGGAGGACTCGATCTGGAG 864

QY 295 LeuLeuAsnGluGlyGlnArgLeuPheSerHisGluGlySer---SerPheGlnMetPhe 313

Db 865 CTTCTCAGCAAGGCACAGCAAGATTTCGAAAAGGCACGTCAGGAGGAGCTTCCAGCTCTTC 924

QY 314 SerSerGluAlaTyrGlyGlnLysAspLeuLeuPheLysAspSerThrSerGluLeuVal 333

Db 925 AGCTCCCTCCTCT---GGGCAAGAACCTGCTTTTCAAGATGGAGGCCCAAGGGTTTTT 981

Alignment Scores:

Pred. No.: 9, 49e-107 Length: 2347

Score: 1304.50 Matches: 320

Percent Similarity: 55.27% Conservative: 110

Best Local Similarity: 41.13% Mismatches: 251

Sequence 2347 BP; 545 A; 655 C; 689 G; 458 T; 0 U; 0 Other;

Claim 3; Page 76-79; 107pp; English.

This sequence encodes the anticoagulant protein, Draculin. Draculin is

isolated from the saliva of vampire bats and is distinct from other known

anticoagulants in that it inhibits both factors IX and X of the blood

coagulation cascade. The anticoagulant activity is not inhibited by

serine protease inhibitors such as PMFS or DFP. The molecular weight of

Draculin ranges from 75-90 kD and may be used in treatment of acute

myocardial infarction, deep vein thrombosis, pulmonary embolism, unstable

angina, transient ischemic attacks, peripheral vascular or bypass

occlusions and disseminated intravascular coagulation. (Updated on 25-MAR

-2003 to correct PN field.)

New anticoagulant protein, draculin, from vampire bat saliva - and

related nucleic acid, vectors, transformed cells and probes, for treating

myocardial infarction, etc.

factor IX; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

acute myocardial infarction; deep vein thrombosis; pulmonary embolism;

unstable angina; transient ischemic attack; peripheral vascular; DFP;

bypass occlusion; disseminated intravascular coagulation; ds.

Desmodus rotundus.

Anticoagulant; Draculin; saliva; vampire bat; inhibition; factor IX;

factor X; blood coagulation cascade; serine protease inhibitor; PMFS;

334 ProfileAlaThrGlnThr---TyrGluAlaTrpLeuGlyHisGluTyrLeuHisAlaMet 352  
 982 AGGATCCCTCAAGGGTGGACGCTGATGTACCTCGGTCCAGCTACCTCACCGTCACTC 1041  
 353 LysGlyLeu-----LeuCysAspProAsnArgLeuProProTyrLeuArgTyrCys 369  
 1042 AAGAACCTGAAGGAATCGGCAGCAGAGGTGGAGGCCCGGGGGCCCGGGTGTGTGTGTGC 1101  
 370 ValLeuSerThrProGluIleGlnLysCysGlyAspMetAlaValAlaPheArgArgGln 389  
 1102 GCGGTGGCCAGAGGAGCTCGCAGTCCAGCAGTGGAGT-----GGC 1146  
 390 ArgLeuLysProGluIleGlnCysValSerAlaLysSerProGlnHisCysMetGluArg 409  
 1147 CAGAGCAATGGCAGCAGTGCCTGCACACAGCCGCTGCACAGAGGAGTGCATCGCCCTG 1206  
 410 IleGlnAlaGluGlnValAspAlaValThrLeuSerGlyGluAspIleTyrThrAlaGly 429  
 1207 GTGCTGAAGAGAGAAGCCGATGCATCAGTCTGGACGGAGGGTTCATCTATATCGCGGC 1266  
 430 LysLysTyrGlyLeuValProAlaAlaGlyGluHisTyrAlaProGluAspSerSerAsn 449  
 1267 ---AAATGTGTTGGCGCTCTCTGCGCAGAGAGCCAAAGATCCGAAGAGGAGCGATTAAC 1323  
 450 -----SerTyrTyrValValAlaValAlaValArgArgAsp 460  
 1324 TTGATTGTGTGAATAGACCACTCGAAGGGTATCGTGTGGCGGTGTGCAGGAAA-- 1380  
 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
 1381 TCAAGTCTGTGCGCTCACCTGGAACCTCTCAGGGGCGACGAAGTCTCTGCCACACCGCTGTG 1440  
 481 GlySerProAlaGlyTyrAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
 1441 GCGAGGACAGCAGGCTGGAACATCCCATGGGTCTGCTC-----TTCAACCAAG 1488  
 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
 1489 ACAGCTCTCTGCAAC-----TTTGATGAATTCCTTCAGTCAAACGTGCGCCCT 1536  
 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
 1537 GGAGCAGACCCGAAAC-----TCCAACCTCTGCGCCCTGTGCGTCGCAATAGCAG 1587  
 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
 1588 GGCACGACAAAGTCGCTCCCAACAGCAACGAGAGGTACTTTCAGCTACGCTGGGGCTTC 1647  
 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
 1648 AGGTGCTGTGCGAAGATGCTGGAGACGTGGCCCTTTGTGAAGCTTCCACTGCTCTGGAG 1707  
 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600  
 1708 AACACGGATGGAAGGGCACCAGAAAGCATGGGCTAAGGATCTCGAAGCTGGAGGACTTTGAG 1767  
 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
 1768 CTCCTGTGCTTGTATGACGCCCGGAAGCCCTGTGTCTGAGTTTGAGACTGCCACCTGGCC 1827  
 621 GlnIleProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
 1828 AGGCCCCGAGTACCGGTGTGGTATCTCGAAAGATAGGTGTACAATACCTGGAGCAGGTG 1887  
 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGly----- 657  
 1888 CTGCTGGACCAA---CAGGCAAGTTTGA-----AGAAATGCAACCCCTCTGCG 1932  
 658 -----PheLysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLys 674  
 1933 CCAGGCAAGTCTCTGTTTCCAGCTGGAAC-----AGAAACCTCTCTGTTCAAT 1983  
 675 AspAlaThrValArgAlaValProValGlyGluLysThrThrTyrArgGlyTyrLeuGly 694

Db	1984	GACAAACCCGAGTGTCTGGCCAAACTCCAGGCAAAACACGATATGAGAAATATTTGGGA	2043
Qy	695	LeuAspTyrValAlaAlaLeuGluGlyMetSerSerClnGlnCysSerGlyAla-----	712
Db	2044	CCAGAGTATGTTCACAGCGGTGCTAATCTG-----AGGCAATGCTCCACCTCCCCACTT	2097
Qy	713	-----AlaAlaProAlaProGlyAla	719
Db	2098	CTGGAAGCCTGTACCTTCCTGAGGAATTGAAACCAAGAAGTGGCCAGCCCTGCCAC	2157
Qy	720	ProLeuLeuProLeuLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAla	737
Db	2158	CCCCACCACCCC-----AAAGCTGCAGCGCCACTGCCCTGGCCCCATCC	2202
RESULT 11			
AAQ12152			
ID	AAQ12152 standard; DNA; 2327 BP.		
XX	AAQ12152;		
XX	25-MAR-2003 (revised)		
DT	24-SEP-1991 (first entry)		
XX	Human transferrin gene.		
DE	iron-binding protein; clone Tf; liver cDNA library; ss.		
XX	Homo sapiens.		
OS			
XX			
FT	Key	Location/Qualifiers	
FT	sig_peptide	31..87	
FT		/*tag= a	
FT	mat_peptide	88..2124	
FT		/*tag= b	
FT		/product= "transferrin"	
XX			
PN	US026651-A.		
XX			
PD	25-JUN-1991.		
XX			
PF	25-APR-1985; 85US-00727335.		
XX			
PR	25-APR-1985; 85US-00727335.		
XX			
PA	(TEXA ) UNIV TEXAS SYSTEM.		
PI	Bowman BH, Yang F;		
XX			
DR	WPI; 1991-207469/28.		
DR	P-ESDB; AAR12499.		
XX			
PT	DNA encoding human transferrin - used for recombinant expression of large		
PT	amts. of human transferrin for use in tissue culture.		
XX			
PS	Claim 14; Table 1; 9pp; English.		
XX			
CC	Clone Tf contg. the entire transferrin gene was isolated after screening		
CC	a human liver cDNA library constructed in the ampicillin resistance		
CC	plasmid pKT218 with a probe pool based on amino acids 309-314 of		
CC	transferrin. The clone has been deposited as ATCC 53106. See also		
CC	AAQ12151. (Updated on 25-MAR-2003 to correct PA field.)		
XX			
SQ	Sequence 2327 BP; 624 A; 535 C; 629 G; 539 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	6.96e-104	Length:	2327
Score:	1272.50	Matches:	313
Percent Similarity:	55.29%	Conservative:	121
Best Local Similarity:	39.87%	Mismatches:	225
Query Match:	32.33%	Indels:	126
DB:	2	Gaps:	34

Alignment Scores:	
Pred. No.:	6.96E-104
Score:	1272.50
Length:	2327
Percent Similarity:	55.2%
Matches:	313
Best Local Similarity:	39.8%
Mismatches:	235
Query Match:	32.3%
Indels:	136
DB:	2
Gaps:	34

US-10-049-957-4 (1-738) x AAQ12152 (1-2327)

1 MetArgGlyProSerGlyAlaLeuTrpLeu-----LeuLeuAlaLeuArgThrValLeu 18  
 Db           |||||           :::           |||           |||           :::           :::  
 31 ATGAGGCTCGCGTGGGAGCCCTGCTGGTCTCGCGCGCTCTGGGGCTGTGCTGGCTGTC 90  
 QY           |||           |||           |||           |||           |||           |||  
 19 GlyGlyMetGluValAlaGlyTrpCysAlaThrSerAspProGluGlnHisCysCysGlyAsn 38  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 91 CCTGATAAACTGTGAGATGGTGTGCACTGTGCGAGCATGAGGCCACTAAGTGC----- 144  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 39 MetSerGluAlaPheArgGlu-----AlaGlyIleGlnProSerLeu 52  
 Db           :::           :::           :::           :::           :::           :::  
 145 -----CAGAGTTCCGGACCATATGAAAAGCGTCATTCATCCGATGGTCCCAGTGT 198  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 53 LeuCysValArgGlyThrSerAlaAspHisCysValGlnLeuIleAlaAlaGlnGluAla 72  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 199 GCTTGTGTGAAGAAAGCCTCCTACCTGATTCATCAGGCCCATTCGCGCAACGAAGCG 258  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 73 AspAlaIleThrLeuAspGlyGlyAlaIleTyrGluAlaGly---LysGluHisGlyLeu 91  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 259 GATGCTGTGACACGTGATGCAAGTGTGTGATGATGCTTACTTGGCTCCCAATAACCTG 318  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 92 LysProValValGlyGluValTyr-----AspGlnGluValGlyThrSerTyrTyrAla 109  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 319 AAGCCTGTGGTGGCAGATTCATTCGGTCAAAAGAGGATCCACAGACTTCTATTATGCT 378  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 110 ValAlaValValArgArgSerSerHisValThrIleAspThrLeuLysGlyValLysSer 129  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 379 GTTCTGTGGTGAAGAGATAGTGGCTTCAGATGAACACAGCTTCGAGGCAAGAAGTCC 438  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 130 CysHisThrGlyIleAsnArgThrValGlyTrpAsnValProValGlyTyrLeuValGlu 149  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 439 TGCACACAGCGTCTAGGCAGGTCCGCTGGGTGGAAACATCCCATAGGCTTACTTTAC--- 495  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 150 SerGlyArgLeuSerValMetGlyCysAsp-----ValLeuLys 162  
 Db           :::           :::           :::           :::           :::           :::  
 496 -----TGTGACTTACCTGAGCCACAGTAAACCTCTCGAGAAA 531  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 163 AlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSer 182  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 532 GCAGTGGCCAATTTCTTCTCGGCAGCTGTGCCCTTGTGGATGGGACGAGCTCC 591  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 183 GluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAspLysSer 202  
 Db           :::           :::           :::           :::           :::           :::  
 592 CAG---CTGTGTCAACTGTGTCCAGG-----TGTGGCTGTGCTCC 627  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 203 ProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGly 222  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 628 ACCCTTAACAAATACTTCGGCTACTCGGAGCGCTTCAAGTGTCTGAAGATGGTGTGG 687  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 223 AspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuPro 242  
 Db           :::           :::           :::           :::           :::           :::  
 688 GATGTGGCTTTGTCAAGCACTTCGACTATATTGAGAACTTGGCAAAACAGGCT----- 741  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 243 SerTrpGlnAlaLeuLeuSerGlnAsp---PheGluLeuLeuCysArgAspGlySer 261  
 Db           :::           :::           :::           :::           :::           :::  
 742 -----GACAGGGACCATGATGAGTGTGCTTTGCTAGTACACACACC 780  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 262 ArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAlaVal 281  
 Db           |||||           |||||           |||||           |||||           |||||           |||||  
 781 CGGAAGCCGGTAGATGAATACAAGACATGCCACTTTGGCCCAAGGCTCTTCTCATACCGTC 840  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 282 ValValArgAlaAspThrAspGlyGly-----LeuIlePheArgLeuLeuAsnGlu 298  
 Db           :::           :::           :::           :::           :::           :::  
 841 GTGGCCGA-----AGTATGGCGGCAAGGAGGACTTGATCTGGAGAGCTTCTCAACCG 894  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 299 GlyGlnArgLeuPheSerHisGluGlySer---SerPheGlnMetPheSerSerGluAla 317  
 Db           :::           :::           :::           :::           :::           :::  
 895 GCCCAGGAACATTTTGGCAAGACAAATCAAAAGAAATTCACACTATTCACTCT---CCT 951  
 QY           |||||           |||||           |||||           |||||           |||||           |||||  
 318 TyrGlyGlnLysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThr 337  
 Db           :::           :::           :::           :::           :::           :::  
 952 CATGGG---AAGGACCTCTGTTTAAGACATCTGCCACGGGTTTTTAAAGATCCCCCA 1008

Qy	338	Gln---ThrTyrGluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeu	356
Db	1009	AGGATGATGCCAAGATGTACCTGGGCTATGATGATCTACTGCATCCCGAATCTACGG	1068
Qy	357	-----CysAspProAsnArgLeuProProTyr	365
Db	1069	GAAGGCACATGCCCAAGAGCCCAACAGATGAATGCAAGCCT	1110
Qy	366	LeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGlyAspMetAlaValAla	385
Db	1111	GTGAAGTGGTGGCTGAGCCACCAGAGAGCTCAAGTGTGATGAGTGGAGTGT---	1167
Qy	386	PheArgGlnArgLeuLysProGluIleGlnCysValSerAlaLysSerProGlnHis	405
Db	1168	-----AACAGTGTAGGGAATAAGAGTGTGTATCAGCAGAGACCAACCGAAGAC	1215
Qy	406	CysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeuSerGlyGluAspIle	425
Db	1216	TGCATGCCCAAGATCATGAATGGAGAAGCTGTATGCCATGAGCTTGGATGAGGAGTTGTCT	1275
Qy	426	TyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGluHisTyrAlaProGlu	445
Db	1276	TACATACGGGC---AAGTGTGCTCTGGTCTCTTGGCAGAAACTACATAAGAGC	1332
Qy	446	AspSerSerAsn-----SerTyrTyrValValAlaValAlaValArgArgAsp	460
Db	1333	GATAATTGTGAGGATACACCAGAGCAGGCTATTCTTCTAGCAGTGGTGAAGAAATCA	1392
Qy	461	SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe	480
Db	1393	GCCTTCTGAC---CTCACCTGGGCAATCTCGAAAGGCAAGAAGTCTCTGCCATACGGCAGTT	1449
Qy	481	GlySerProAlaGlyTyrAspValProValGlyValAlaLeuIleGlnArgGlyPheIleArg	500
Db	1450	GGCAGAACCCCTGGCTGGAGACATCCCATCGGCTGCTC-----TACAATAAG	1497
Qy	501	ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro	520
Db	1498	ATCAACCACTGC-----AGATTTTGATGAATTTTTTCAGTGAAGGTGTGTGCCCT	1545
Qy	521	ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluCln	540
Db	1546	GGGTCTAAGAAA-----GACTCCAGTCTCTGTGAAGCTGTGTATGGGC-----TCA	1590
Qy	541	GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe	560
Db	1591	GGCCTAAACCTGTGTGNAACCAACACAGAGGGATACTACGGCTACAGAGCGCTTTC	1650
Qy	561	ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp	580
Db	1651	AGGTGCTGTGTTGAGAAG---GGAGATGTGGCCTTTGTGAAACACCAGACTGCTCCACAG	1707
Qy	581	AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu	600
Db	1708	AACACTGGGGGAAAAAACCTCGATCCATGGGCTAAGAATCTGAATGAAAAGACTATGAG	1767
Qy	601	LeuLeuCysProAsnGlyAlaArgAlaGluValSerClnPheAlaAlaCysAsnLeuAla	620
Db	1768	TTGCTGTGCTTATGATGTATCCAGAAACCTGTGGAGGAGTATGCAACTGCCACCTGGCC	1827
Qy	621	GlnIleProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly	640
Db	1828	AGACCCCGAATCACGCTGTGGTTCACACGGAAGATAAGGAAGCT---TCGCTCCACAG	1884
Qy	641	LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsn-----LysAsnGlyPhe	658
Db	1885	ATATTACGTCAACAGCAGCAGCACCTATTTTGGAAGCAACGTAACCTGCTCGGCAACTTT	1944
Qy	659	LysMetPheAspSerSerAsnTyrHisGlyClnAspLeuLeuPheLysAspAlaThrVal	678
Db	1945	TGTTTGTTCGGTTCGGAACCC-----AAGACCTCTGTTTCAGAGATGACAGTAGTA	1995

db  
952 CATGGG-- --AAGGACCTGCTGTTTAAGGACTCTGCCCA CGGGTTTTTAAAGTCCCCCA 1008

QY 679 ArgAlaValProValGluGlyThrThrThrArgGlyTrpLeuGlyLeuAspTyrVal 698  
 Db 1996 TGTGGCCAACTTCATGACAGAACACATATGAAAATACTTAGGAGAAATATGTC 2055  
 QY 699 AlaAlaLeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGly 718  
 Db 2056 AAGGCTGTTGGTAACCTG-----AGAAATGCTCCAACTCATCA----- 2094  
 QY 719 AlaProLeuProLeuLeu-LeuProAlaLeuAlaAlaArgLeu----- 733  
 Db 2095 -----CTCTGGAAGCCTGCACTTTCCGAGACCTTAAATCTCAGAGT 2139  
 QY 734 ----LeuProPro 736  
 Db 2140 AGGGCTGCCACCA 2152

## RESULT 12

ABZ82283  
 ID ABZ82283 standard; cDNA; 2318 BP.

XX AC ABZ82283;

XX DT 11-AUG-2003 (first entry)

XX DE Human transferrin coding sequence.

XX KW Human; transferrin; neuroprotective; cerebroprotective; vasotrophic;  
 KW antiparkinsonian; nootropic; anti-HIV; antisthmatic; antiallergic;  
 KW cytostatic; immunosuppressive; antiatherosclerotic; cardiant;  
 KW gynaecological; immunostimulant; antianaemic; haemostatic;  
 KW antiinflammatory; dermatological; antibacterial; virucide; antiparasitic;  
 KW fungicide; hepatotropic; antirheumatic; antiarthritic; antigout;  
 KW tranquilizer; vulnerary; antidiabetic; nephrotropic; antipyretic;  
 KW gastrointestinal; gene therapy; transgenic animal; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 51..2150

FT /tag= a

FT /product= "Human transferrin"

FT sig\_peptide 51..107

FT /tag= b

FT mat\_peptide 108..2147

FT /tag= c

XX WO2003020746-A1.

XX PD 13-MAR-2003.

XX PF 30-AUG-2002; 2002WO-US027637.

XX PR 30-AUG-2001; 2001US-03157459P.

XX PR 30-NOV-2001; 2001US-0334059P.

XX PA (BIOR-) BIOREXIS PHARM CORP.

XX PI Prior CP;

XX WPI; 2003-332916/31.

XX DR P-PSDB; ABP72819.

XX PT New fusion protein, useful in the diagnosis and treatment of diseases or  
 PT disorders relating to the respiratory, cardiovascular and digestive  
 PT systems, comprises a transferrin protein fused to a therapeutic protein.  
 XX Disclosure; Page 277-279; 298pp; English.

PS The present sequence is that of a nucleic acid encoding human transferrin  
 CC (Tf). The invention relates to modified Tf fusion proteins comprising at  
 CC least one therapeutic protein, polypeptide or peptide, in which the Tf  
 CC portion is engineered to extend the serum half-life or bioavailability of  
 CC the molecule. The modified Tf fusion protein preferably comprises a human

CC Tf moiety that has been modified to reduce or prevent glycosylation, iron  
 CC binding and/or transferrin receptor binding. Nucleic acids encoding such  
 CC fusion proteins, vectors, host cells and transgenic animals which produce  
 CC the fusion protein in their serum or milk are also claimed. The modified  
 CC fusion protein is useful for treating a disease or disease symptom, or  
 CC for delivering a therapeutic agent complexed to the ferric iron of  
 CC transferrin to the inside of a cell or across the blood-brain barrier.  
 CC The modified fusion protein, or a nucleic acid encoding it, can be used  
 CC in the diagnosis, prognosis, prevention and/or treatment of diseases  
 CC and/or disorders of the endocrine, nervous, immune, respiratory,  
 CC cardiovascular, reproductive and digestive systems, diseases and/or  
 CC disorders relating to the blood or to cell proliferation, inflammatory  
 CC conditions, and to treat viral, fungal, bacterial or parasitic infection  
 XX  
 SQ Sequence 2318 BP; 598 A; 545 C; 636 G; 539 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.05e-103 Length: 2318  
 Score: 1270.50 Matches: 313  
 Percent Similarity: 55.29% Conservative: 121  
 Best Local Similarity: 39.87% Mismatches: 225  
 Query Match: 32.28% Indels: 126  
 DB: 7 Gaps: 34

US-10-049-957-4 (1-738) x ABZ82283 (1-2318)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeu-----LeuLeuAlaLeuArgThrValLeu 18  
 Db 51 ATGAGGCTCGCGCGTGGGAGCCCTGCTGCTCGCCCGCTCTGGGCTGTGTGCTGTC 110  
 QY 19 GlyGlyMetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsn 38  
 Db 111 CCTGATAAACTGTGAGATGTGTGCTGCTGCGAGCATGAGGCCACTAAGTGC----- 164  
 QY 39 MetSerGluAlaPheArgGlu-----AlaGlyIleGlnProSerLeu 52  
 Db 165 -----CAGAGTTCGCGGACCACATAGAAAGCGTCATTCATCCATCGATGGTCCCAAGTGT 218  
 QY 53 LeuCysValArgGlyThrSerAlaAspHisCysValGlnLeuIleAlaAlaGlnGluAla 72  
 Db 219 GCTTGTGTGAAGAAAGCCCTCTCTACCTTGTATTCATCAGGGCCATTGGCGCAACGAAACG 278  
 QY 73 AspAlaIleThrLeuAspGlyGlyAlaIleTyrGluAlaGly---LysGluHisGlyLeu 91  
 Db 279 GATGCTGTGACACTGGATCGAGGTTGTGTATGATGCTTACCTGGCTCCCAATAACCTG 338  
 QY 92 LysProValValGlyGluValTyr-----AspGlnGluValGlyThrSerTyrTyrAla 109  
 Db 339 AAGCTGTGTGGCAGAGTTCTATGGGTCAAAAGAGGATCCACAGACTTTCTATTATGCT 398  
 QY 110 ValAlaValValArgArgSerSerHisValThrIleAspThrLeuLysGlyValLysSer 129  
 Db 399 GTTGTGTGTGAAGAGGATAGTGGCTTCAGATGAACACAGCTTCGAGGGAAGAAGTCC 458  
 QY 130 CysHisThrGlyIleAsnArgThrValGlyTrpAsnValProValGlyTyrLeuValGlu 149  
 Db 459 TCCACACGGGCTAGGACAGGTCCGCTGGGTGGAAACATCCCATAGGCTTACTTTAC--- 515  
 QY 150 SerGlyArgLeuSerValMetGlyCysAsp-----ValLeuLys 162  
 Db 516 -----TGTCATTTACTGTAGCCACCGTAAACCTCTTGAGAAA 551  
 QY 163 AlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSer 182  
 Db 552 GCAGTGGCCAAATTCCTTCCTCGGCGAGCTGTGCCCTTGTGCGGATGGACGAGCTTCCCC 611  
 QY 183 GluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGlyGluValCysAspLysSer 202  
 Db 612 CAG---CTGTGTCAACTGTGTCCAGGG-----TGTTGGTGTGCTCC 647  
 QY 203 ProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGly 222  
 Db 648 ACCCTTAACCAATACTTCGGCTACTCGGGAGCCCTTCAAGTGTCTCTGAAGGATGTTGCTGGG 707

```
QY 223 AspValAlaPheValIlyHisSerThrValLeuGluAsnThrAspGlyLysThrLeuPro 242
DB 708 GATGTGGCTTTGTCAGACCTGACTATATTGGAGAACTTGGCAACAGGCT-----761
QY 243 SerTrpGlyClnAlaLeuLeuSerGlnAsp-----PheGluLeuLeuCysArgAspGlySer 261
DB 762 -----GACAGGACCAATATGAGCTGCTTTGGCTGGACACAC 800
QY 262 ArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAlaVal 281
DB 801 CGGAACCGGTAGATGAATACAGGACTGCCCTTGGCCAGGTCCCTTCTCATACCGTC 860
QY 282 ValValArgAlaAspThrAspGlyGly-----LeuIlePheArgLeuLeuAsnGlu 298
DB 861 GTGGCCCGA-----AGTATGGCGCGCAAGGAGGACTTGATCTGGGAGCTTCTCAACACAG 914
QY 299 GlyGlnArgLeuPheSerHisGluGlySer-----SerPheClnMetPheSerSerGluAla 317
DB 915 GCCCAGGAACATTTTGGCAAGACAAATCAAAAGAAATTCCAACTATTACAGCTCT---CCT 971
QY 318 TyrGlyGlnIlyAspLeuPheLysAspSerThrSerClnLeuValProIleAlaThr 337
DB 972 CATGGG---AAGGACCTGCTTTAAGGACTCTGCCACGGGTTTTTAAAGTCCCCCC 1028
QY 338 Gln---ThrTyrGluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeu 356
DB 1029 AGGATGGATGCCAAGATGATACCTGGGCTATGAGTATGTCACTGCCATCCGGAATCTACGG 1088
QY 357 -----CysAspProAsnArgLeuProTyr 365
DB 1089 GAAGGCACATGCCCAGAGCCCAACACAGATGAATGCAAGCCT-----1130
QY 366 LeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGlyAspMetAlaValAla 385
DB 1131 GTGAAGTGGTGGCTGAGCCACACAGAGAGGCTCAAGTGTGATGAGTGGAGTGT---1187
QY 386 PheArgArgGlnArgLeuLysProGluIleGlnCysValSerAlaLysSerProGlnHis 405
DB 1188 -----NACAGTGTAGGGAATAAGATGTGTATCAGCAGACCCAGAGAC 1235
QY 406 CysMetGluArgIleGlnAlaGluInValAspAlaValThrLeuSerGlyGluAspIle 425
DB 1236 TGCATGCCAAGATCATGAATGAGAGAGCTGTATGCCATGAGCTTGGATGGAGGTTTGTG 1295
QY 426 TyrThrAlaGlyLysIlySerGlyLeuValProAlaAlaGlyClnHisTyrAlaProGlu 445
DB 1296 TACATAGCGGGC---AAGTGTGTGTGGTGCCTCTTGGCAGAAACACTACAATAAGAGC 1352
QY 446 AspSerSerAsn-----SerTyrTyrValValAlaValValArgArgAsp 460
DB 1353 GATAATTGTGAGGATACACAGAGCGGAGGTATTTGCTGTAGCAGTGGTGAAGAATCA 1412
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480
DB 1413 GCTTCTGAC---CTCACCTGGGCAATCTGAAAGGCAAGAAGTCTGCCATACGGCAGTT 1469
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500
DB 1470 GGCAGAACCCCTGGCTGGAAACATCCCATCGGCTGGCTCTC-----TACAATAAG 1517
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520
DB 1518 ATCAACCATGCTC-----AGATTTGATGAATTTTTCAGTGAAGTGTGTCCTT 1565
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540
DB 1566 GGGTCTAAGAAA-----GACTCCAGTCTCTGTAAGCTGTGTATGGGC-----TCA 1610
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTrpGlyTyrArgGlyAlaPhe 560
DB 1611 GGCCTAAACCTGTGTGAACCCCAACAAAGAGGGGATCTACTACGGCTACACAGCGCTTTC 1670
```

---

```
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580
DB 1671 AGGTGTCTGTTTGAAG---GGAGATGTGGCTTTGTGAAACACCCAGACTGTCCACAG 1727
QY 581 AsnThrAsnGlyHisAsnSerClnProTrpAlaAlaGluLeuArgSerClnAspTyrGlu 600
DB 1728 AACACTGGGGGAAAAAACCCCTGATCCATGGGCTAAGAATCTGAATGAAAGAACTATGAG 1787
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
DB 1788 TTGCTGTGCTTGTGTTGACAGGAAACCTGTGGAGGAGTATGCCAACTGCCACCTGGCC 1847
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
DB 1848 AGAGCCCGGAATCAGCTGTGTGCACACGGAAGATAAGAAAGCT---TGCCTCCACAA 1904
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsn-----LysAsnGlyPhe 658
DB 1905 ATATTACGTCAACAGCAGCACCCTATTGGAAGCAACGTAACCTGCTCGGGCAACTTT 1964
QY 659 LysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrVal 678
DB 1965 TGTTTGTTCGGTCCGAAACC-----AAGGACCTTCTGTTCCAGATGCACAGAT 2015
QY 679 ArgAlaValProValGlyGluLysThrTyrArgGlyTrpLeuGlyLeuAspTyrVal 698
DB 2016 TGTTGGCCAAACTTCATGACAGAAACACATATGAAAAATCTTAGGAGAAGATATGTC 2075
QY 699 AlaAlaLeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGly 718
DB 2076 AAGGCTGTGTGTAACCTG-----AGAAATGCTCCACTCATCA-----2114
QY 719 AlaProLeuLeuProLeuLeu-LeuProAlaLeuAlaAlaArgLeu-----733
DB 2115 -----CTCTGGAAGCTGCACCTTTCCGTAGACCTTAAATCTCAGAGT 2159
QY 734 ----LeuProPro 736
DB 2160 AGGGCTGCCACCA 2172
RESULT 13
AAZ24196
ID AAZ24196 standard; cDNA; 2327 BP.
XX
AC AAZ24196;
XX
DT 04-FEB-2000 (first entry)
XX
DE Human serum transferrin cDNA.
XX
KW Serum transferrin; human; mutant; metal binding; iron chelator;
KW metal chelation therapy; toxic metal; metal overload; thalassemia; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..2127
FT /tag= b
FT /product= "serum transferrin"
FT sig_peptide 1..87
FT mat_peptide 88..2124
FT /tag= a
FT /tag= c
XX
PN US986067-A.
XX
PD 16-NOV-1999.
XX
PF 28-DEC-1993; 93US-00175158.
XX
PR 08-FEB-1991; 91US-00652869.
XX 06-FEB-1992; 92US-00832029.
XX
```

(UYVE-) UNIV VERMONT & STATE AGRIC COLLEGE.  
(UYBR-) UNIV BRITISH COLUMBIA.

PA Funk WD, Macgillivray RTA, Woodworth RC, Mason AB;  
XX P-PSDB; AAY50717.

XX WPI; 2000-022329/02.

XX P-PSDB; AAY50717.

XX Recombinant human serum transferrin mutants with altered metal binding  
PT properties are useful in metal chelation therapy and to clear excess  
PT toxic metals in patients.

XX Disclosure; Col 19-24; 26pp; English.

XX This invention describes novel recombinant human serum transferrin  
CC mutants with altered metal binding properties. The products of the  
CC invention act as iron chelators. The recombinant transferrins can be used  
CC in metal chelation therapy to bind and clear excess toxic metals in  
CC patients suffering from metal overloads. In particular transferrin  
CC mutants which bind iron with higher avidity than natural transferrin can  
CC be administered to individuals suffering from thalassemia to remove  
CC excess toxic iron from the body. Half-molecules or transferrin mutants  
CC with altered metal ion selectivities could be used to clear other toxic  
CC metals e.g. lead, mercury, cadmium copper or zinc from the body.  
CC Recombinant full length transferrin can also be used in non serum  
CC supplements or in tissue culture media. A transferrin half-molecule is  
CC advantageous as it is able to pass through the glomeruli of the kidney  
CC and can be excreted in the urine, unlike the whole-proteins, so that the  
CC metal is not only chelated but also cleared from the body. The single  
CC half-molecules do not bind to transferrin receptors on the membrane of  
CC tissue cells and therefore do not deliver the iron to these tissues. A  
CC further advantage is that the human body probably recognizes the half-  
CC molecules as 'self' and would not elicit an immunological response. Using  
CC recombinant transferrin avoids the risk of contamination with HIV or  
CC hepatitis virus associated with transferrin purified from human serum.  
CC This sequence encodes the human serum transferrin described in the  
CC invention

XX Sequence 2327 BP; 623 A; 538 C; 627 G; 539 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 1.05e-103 Length: 2327  
Score: 1270.50 Matches: 313  
Percent Similarity: 55.29% Conservative: 121  
Best Local Similarity: 39.87% Mismatches: 225  
Query Match: 32.28% Indels: 126  
DB: 3 Gaps: 34

US-10-049-957-4 (1-738) x AA224196 (1-2327)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeu-----LeuLeuAlaLeuArgThrValLeu 18  
DB ATGAGGCTCGCGTGGAGCCCTGCTGGTCTGGCGGCTCTGGGCTGTGTCTGGCTGTC 90  
QY 19 GlyGlyMetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsn 38  
DB CCTGATAAACTGTGAGATGGTGTGACGTGTCGAGTGTGGAGCATGAGCCACTAAGTGC----- 144  
QY 39 MetSerGluAlaPheArgGlu-----AlaGlyIleGlnProSerLeu 52  
DB :-----CAGATTTCGCGCACCATTATGAAAGCGTCAITTCATCGATGTCCTCCAGTGT 198  
QY 53 LeuCyValArgGlyThrSerAlaAspHisCysValGlnLeuLeuAlaGlnGluAla 72  
DB GCTTGTGTGAAGAACCTCTTACCTTGATTGATTCAGGGCCATTTCGCGCAACCAAGCG 258  
QY 73 AspAlaIleThrLeuAspGlyGlyAlaIleTyrGluAlaGly---LysGluHisGlyLeu 91  
DB :-----GATGCTGTGACATGAGTGTGGTCTATGATGCTTACTTGGCTCCCAATACCTG 318  
QY 92 LysProValValGlyGluValTyr-----AspGlnGluValGlyThrSerTyrAla 109  
DB :-----AAGCCTGTGTGGCAGAGTCTTATGGGTCAAAAGAGGATCCACAGACTTCTTATTATGCT 378

QY 110 ValAlaValValArgArgSerSerHisValThrIleAspThrLeuLysGlyValLysSer 129  
DB GTTGCTGTGGTGAAGAGGATAGTGGCTTCAGATGAACACCGAGCTTCAGGGAAGAGTCC 438  
QY 130 CysHisThrGlyIleAsnArgThrValGlyTyrAsnValProValGlyTyrLeuValGlu 149  
DB TGGCACCGGTCTAGGACGGTCCGCTGGGTGGAAACATCCCATAGGCTTACTTTTAC--- 495  
QY 150 SerGlyArgLeuSerValMetGlyCysAsp-----ValLeuLys 162  
DB :-----TGTCACTTACCTGAGCCACGCTAAACCTCTTGGAGAA 531  
QY 163 AlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSer 182  
DB :-----GAGTGGGCAATTTCTTCGCGGAGCTGTGCCCTTGTGGATGGGACGACTTCCC 591  
QY 183 GluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAspLysSer 202  
DB :-----CTGTGTCAACTGTGTCCAGG-----TGTGGCTGCTCC 627  
QY 203 ProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGly 222  
DB ACCCTTAACCAATACTTCGGCTACTTCGGGAGCTTCAAGTGTCTGAGAGGATGCTGGG 687  
QY 223 AspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuPro 242  
DB GATGTGGCTTGTCAAGCACTCGACTATATTTGAGAACTTGGCAACAGAGCT----- 741  
QY 243 SerTrpGlyGlnAlaLeuLeuSerGlnAsp---PheGluLeuLeuCysArgAspGlySer 261  
DB :-----GACAGGACCCAGTATGAGCTGCTTTGCTAGACAACACC 780  
QY 262 ArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAlaVal 281  
DB :-----CGGAGCGGTAGATGAATACAGGACTGTCACCTTGGCCAGGTCCTTCTCATAC 840  
QY 282 ValValArgAlaAspThrAspGlyGly-----LeuIlePheArgLeuLeuAsnGlu 298  
DB :-----AGTATGGCGGCAAGAGAGACTTGTATCTGGGAGCTTCTCAACCC 894  
QY 299 GlyGlnArgLeuPheSerHisGluGlySer---SerPheGlnMetPheSerSerGluAla 317  
DB :-----GCCAGGAACTTTTGGCAAGACAAATCAAAAGAAATTTCCAACATTTTCAAGCTCT---CCT 951  
QY 318 TyrGlyGlnLysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThr 337  
DB :-----CATGG---AAGGACCTGCTGTTTAAGGACTCTGCCACCGGGTTTTAAAAAGTCCCCCA 1008  
QY 338 Gln---ThrTyrGluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeu 356  
DB :-----AGGATGGATGCCAAGATGTACTCTGGCTATGATGTATGTCACTGCCATCCGAATCTACGG 1068  
QY 357 :-----CysAspProAsnArgLeuProTyr 365  
DB :-----GAAAGCATGCCAGAAAGCCCAACAGATCAATGCAAGCCT----- 1110  
QY 366 LeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGlyAspMetAlaValAla 385  
DB :-----GTGAAGTGTGTGCTGAGCCACACGAGAGGCTCAAGTGTGATGATGAGTGAGTGT--- 1167  
QY 386 PheArgArgGlnArgLeuLysProGluIleGlnCysValSerAlaLysSerProGlnHis 405  
DB :-----AACAGTGTAGGAAATAGATGTGTATCAGCAGAGACCCACCGAAGAC 1215  
QY 406 CysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeuSerGlyGluAspIle 425  
DB :-----TGCATGCCAAGATCATGAATGGAGAGCTCATGTCATGAGCTTGTGATGGAGGTTTGTCTC 1275  
QY 426 TyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGluHisTyrAlaProGlu 445  
DB :-----TACATAGCGGC---AAGTGTGGTCTGTGCTCTTGGCAGAAACTCAATAAGAGC 1332

QY 446 AspSerSerAsn-----SerTyrTyrValValAlaValValArgArgAsp 460  
 Db 1333 GATAATTGTGAGGATACACAGAGCGAGGTATTTTCTGTACGATGGTGAAGAATCA 1392  
 QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
 Db 1393 GCTTCTGAC---CTCACCTGGGACATCTGAAAGGCAAGAGTCTGCCATACGGCAGTT 1449  
 QY 481 GlySerProAlaGlyTyrAspValProValGlyAlaLeuLeuGlnArgGlyPheIleArg 500  
 Db 1450 GGCAGAACCGCTGGCTGGAACATCCCATGGCTGCTC-----TACAATAAG 1497  
 QY 501 ProlYsnPysCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
 Db 1498 ATCAACACCTGC-----AGATTTCATGATCAATTTTTCATGGAAGTTTGTGCCCT 1545  
 QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
 Db 1546 GGGTCTAAGAA-----GACTCCAGTCTCTGTAAGCTGTGTATGGC-----TCA 1590  
 QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
 Db 1591 GGCCTAAACCTGTGTGAACCAACAAAGAGGGATACCTACGGCTACACAGCGCTTTC 1650  
 QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
 Db 1651 AGGTGTCTGTGTGAGAG---CGAGATGTGGCTTTGTGAACACACAGACTGTCACACAG 1707  
 QY 581 AsnThrAsnGlyHisAsnSerGluProTTPAlaAlaGluLeuArgSerGluAspTyrGlu 600  
 Db 1708 AACACTGGGGGAAAAACCTGATCCATGGCTAAGATCTGAATGAAAAGACTATGAG 1767  
 QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
 Db 1768 TTGCTGTGCTTGTGTACAGGAAACCTGTGGAGGAGTATCGGAACCTGCCACCTGGCC 1827  
 QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
 Db 1828 AGAGCCCGGAATCAGCTGTGGTCACACGGAAAGATAAGGAAGCT---TGCCTCCCAAG 1884  
 QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsn-----LysAsnGlyPhe 658  
 Db 1885 ATATTAGTCAACAGACGACCTATTGTGGAGCAAGTAACTGACTGCTCGGGCAACTTT 1944  
 QY 659 LysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrVal 678  
 Db 1945 TGTTTGTTCGGTCGGAACCC-----AAGGACCTTCTGTTCAGATGACACAGTA 1995  
 QY 679 ArgAlaValProValGlyLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrVal 698  
 Db 1996 TGTTTGGCCAACTTTCATGACAGAAACACATATGAAAAATACTTAGGAGAAGTAATGTC 2055  
 QY 699 AlaAlaLeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGly 718  
 Db 2056 AAGCGTGTGGTAACCTG-----AGAAATGCTCCACCTCATCA----- 2094  
 QY 719 AlaProLeuLeuProLeuLeuLeuProAlaLeuAlaAlaArgLeu----- 733  
 Db 2095 -----CTCCTGGGAAGCTGCATCTTCGGTAGACTTAAATCTCAGAGCT 2139  
 QY 734 -----LeuProPro 736  
 Db 2140 AGGGCTGCCACCA 2152  
 RESULT 14  
 ABK84172  
 ID ABK84172 standard; cDNA; 2347 BP.  
 XX  
 AC ABK84172;  
 XX  
 DT 14-AUG-2002 (first entry)  
 XX  
 DE Human cDNA differentially expressed in granulocytic cells #743.

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;  
 KW viral infection; parasitic infection; protozoal infection;  
 KW fungal infection; sterile inflammatory disease; psoriasis;  
 KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
 KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
 KW adult respiratory distress syndrome; inflammatory bowel disease;  
 KW Crohn's disease; ulcerative colitis; periodontal disease;  
 KW granulocyte activation; chronic inflammation; allergy.  
 XX Homo sapiens.  
 XX WO200228999-A2.  
 XX 11-APR-2002.  
 XX 03-OCT-2001; 2001WO-US030821.  
 XX 03-OCT-2000; 2000US-0237189P.  
 XX (GENE-) GENE LOGIC INC.  
 XX Bearzer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
 XX WPI; 2002-435328/46.  
 XX Detecting granulocyte activation by detecting differential expression of  
 genes associated with granulocyte activation, which serves as diagnostic  
 markers that is useful for monitoring disease states and drug toxicity.  
 XX Claim 1; SEQ ID NO 743; 114pp; English.  
 XX The invention relates to detecting (M1) granulocyte (GC) activation  
 (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
 DNA chip analysis as given in the specification, and comparing the  
 expression level to an expression level in an unactivated GC, where  
 differential expression of Gs is indicative of GCA. Also included are  
 modulating (M2) GA by contacting GC with an agent that alters the  
 expression of at least one gene in Gs; (2) screening (M3) for an agent  
 capable of modulating GCA or an inflammation (especially chronic) in a  
 tissue, an allergic response in a subject, exposure of a subject to a  
 pathogen or sterile inflammatory disease using the gene expression  
 profile; (3) detecting (M4) an inflammation (especially chronic) in a  
 tissue, an allergic response in a subject, exposure of a subject to a  
 pathogen or sterile inflammatory disease, by detecting the level of  
 expression in a sample of the tissue of gene(s) from Gs, where the level  
 of expression of the gene is indicative of inflammation; (4) treating  
 (M5) an inflammation (especially chronic) or in a tissue, an allergic  
 response in a subject, exposure of a subject to a pathogen or sterile  
 inflammatory disease, by contacting a tissue having inflammation with an  
 agent that modulates the expression of gene(s) from Gs in the tissue. M1  
 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful  
 for screening an agent capable of modulating GCA preferably in an  
 inflammation in a tissue; M4 is useful for detecting an inflammation  
 (especially chronic) in a tissue, an allergic response in a subject,  
 exposure of a subject to a pathogen or sterile inflammatory disease (e.g.  
 psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,  
 cardiac reperfusion injury, renal reperfusion injury, ARDS, adult  
 respiratory distress syndrome, inflammatory bowel disease, Crohn's  
 disease, ulcerative colitis, periodontal disease; also bacterial  
 infection, viral infection, parasitic infection, protozoal infection,  
 fungal infection and M5 is useful for treating one of the above  
 conditions. The present sequence represents a gene differentially  
 expressed in granulocytes. Note: The sequence data for this patent did  
 not form part of the printed specification, but was obtained in  
 electronic format directly from WIPO at  
 ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 2347 BP; 604 A; 555 C; 647 G; 540 T; 0 U; 1 Other;

Alignment Scores: 1.06e-103 Length: 2347  
 Pred. No.: 1270.50 Matches: 313  
 Score:





```
QY 659 LysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrVal 678
Db 1994 TGTTTGTTCGGTCGGAACC-----AAGGACCTTCTGTTTCAGAGTGCACAGTA 2044
QY 679 ArgAlaValProValGlyGluLysThrThrTyrArgGlyTrpLeuGlyLeuAspTyrVal 698
Db 2045 TGTTTGGCCAACTTCATGACGAACACATATGAAAAATACTTAGGAGAAGAAATATGTC 2104
QY 699 AlaAlaLeuGluGlyMetSerSerGlnGlnCysSerGlyValAlaAlaProAlaProGly 718
Db 2105 AAGGCTGTTCGTTAACCTG-----AGAAAATGCTCCACCTCATCA----- 2143
QY 719 AlaProLeuLeuProLeuLeu-LeuProAlaLeuAlaAlaArgLeu----- 733
Db 2144 -----CTCCTGGAAGCCTGCACCTTCGTTAGACCTTAAATAATCTCAGAGGT 2188
QY 734 -----LeuProPro 736
Db 2189 AGGCGTGCCACCA 2201
RESULT 15
ABN96543
ID ABN96543 standard; DNA; 2347 BP.
AC ABN96543;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3041 used to diagnose liver cancer.
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumor; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX Homo sapiens.
OS
PN WC200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030589.
XX
PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX
PS Claim 1; SEQ ID NO 3041; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
```

SQ Sequence 2347 BP; 604 A; 555 C; 647 G; 540 T; 0 U; 1 Other;

## Alignment Scores:

Pred. No.:	1-06e-103	Length:	2347
Score:	1270.50	Matches:	313
Percent Similarity:	55.29%	Conservative:	121
Best Local Similarity:	39.87%	Mismatches:	225
Query Match:	32.28%	Indels:	126
DB:	6	Gaps:	34

US-10-049-957-4 (1-738) x ABN96543 (1-2347)

```
QY 1 MetArgGlyProSerGlyAlaLeuTrpLeu-----LeuLeuAlaLeuArgThrValLeu 18
Db 80 ATGAGGCTCGCGTGGAGCCCTGCTGTGCTCGCGCTCTGCGGCTGTGTCTGGGTGTC 139
QY 19 GlyGlyMetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsn 38
Db 140 CCTGATAAAACCTGTGAGATGGTGTGCAGTGTCCGAGCATGAGGCCACTTAAGTGC----- 193
QY 39 MetSerGluAlaPheArgGlu-----AlaGlyIleGlnProSerLeu 52
Db 194 -----CAGAGTTTCGCGACCATATGAAAAGCGTCATTCATCCGATCGTCCAGTGT 247
QY 53 LeuCysValArgGlyThrSerAlaAspHisCysValGlnLeuLeuAlaGlnGluAla 72
Db 248 GCTTGTGTGAAGAAAGCCTCTACCTTGATTGTCATCAGGGCCATTCGCGCAACAGACG 307
QY 73 AspAlaIleThrLeuAspGlyAlaIleTyrGluAlaGly---LysGluHisGlyLeu 91
Db 308 GATGCTGTGACACTGCGATGCGAGTTTGGTGTATGATGCTTACTCGCTCCCAATAACCTG 367
QY 92 LysProValValGlyGluValTyr-----AspGlnGluValGlyThrSerTyrTyrAla 109
Db 368 AAGCCTGTGTGGCAGAGTTCATGGGTCAAAGAGGATCCACAGACTTCTATTATGCT 427
QY 110 ValAlaValValArgSerSerHisValThrIleAspThrLeuLysGlyValLysSer 129
Db 428 GTTGTGTGTGAAGAGGATAGTGGCTTCCAGATGAACAGCTTCGAGGCAAGAGTCC 487
QY 130 CysHisThrGlyIleAsnArgThrValGlyTyrAsnValProValGlyTyrLeuValGlu 149
Db 488 TGCCACACGGGTCTAGGCAGGTCCGCTGGGTGGAGACATCCCATAGGCTTACTTTAC--- 544
QY 150 SerGlyArgLeuSerValMetGlyCysAsp-----ValLeuLys 162
Db 545 -----TGTGACTTACCTGAGCCACGTAACACTCTTGAGAAA 580
QY 163 AlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSer 182
Db 581 CGAGTGGCCAAATTTCTTCTCGGGCAGCTGTGCCCTTGTGCGATGGGACGACTTCCCC 640
QY 183 GluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAspLysSer 202
Db 641 CAG---CTGTGCAACTGTGTCCAGGG-----TGTGGCTGCTCC 676
QY 203 ProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGlyAlaGly 222
Db 677 ACCCTTAACCAATCTTCGGCTACTCAGGAGCGCTTCAAGTGTCTGAAGATGGTGTGGG 736
QY 223 AspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuPro 242
Db 737 GATGTGGCCTTTGTCAAGCACTCGACTATATTGGAAGCTTGGCAACAAGGCT----- 790
QY 243 SerTrpGlyGlnAlaLeuLeuSerGlnAsp-----PheGluLeuLeuCysArgAspGlySer 261
Db 791 -----GACAGGGACCATATGAGCTGTCTTGGCTGGCAACACACC 829
QY 262 ArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAlaVal 281
Db 830 CGGAAGCCGGTAGATGAATACAAAGGACTCGCCACTTGGCCCGAGGTCTCTCTACACCGTC 889
QY 282 ValValArgAlaAspThrAspGlyGly-----LeuIlePheArgLeuLeuAsnGlu 298
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 16, 2004, 19:33:51 ; Search time 157 Seconds  
(without alignments)  
2608.623 Million cell updates/sec

Title: US-10-049-957-4

Perfect score: 3936

Sequence: 1 MRGPGCALWLLALRTVLGG.....APFLPLLLPALAARLLPAL 738

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DRV=xlh  
-Q=/cgn2\_1/USFTO\_spool/US10049957/runat\_14052004\_100250\_1281/app\_query.fasta\_1.903  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10049957 @CNC 1.1 56 @runat\_14052004\_100250\_1281 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3936	100.0	2368	2	US-08-520-933-1
2	3936	100.0	2368	4	US-09-285-040-1
3	3936	100.0	2368	6	5262177-1
4	1270.5	32.3	2327	2	US-08-175-158A-1
5	1269.5	32.3	2037	3	US-08-848-760B-25
6	1263	32.1	2360	1	US-08-145-681-1
7	1263	32.1	2360	1	US-08-250-308-1
8	1263	32.1	2360	1	US-08-453-703-1
9	1263	32.1	2360	2	US-08-456-106-1
10	1263	32.1	2360	3	US-08-456-108-1
11	1263	32.1	2360	3	US-09-265-577-1
12	1263	32.1	2360	4	US-09-633-739-1

13	1263	32.1	2360	5	PCT-US93-03614-1
14	1261	32.0	2117	2	US-08-655-640-1
15	1260.5	32.0	2124	2	US-08-655-640-3
16	1257.5	31.9	3427	1	US-08-470-058-1
17	1257.5	31.9	3427	3	US-09-037-188-1
18	1257.5	31.9	3427	3	US-09-285-310-1
19	1256	31.9	2619	4	US-09-377-497-14
20	1252	31.8	2086	3	US-08-724-586-1
21	1252	31.8	2086	4	US-09-421-632-1
22	1252	31.8	2086	4	US-09-932-190-1
23	1250.5	31.8	2347	1	US-08-145-681-3
24	1250.5	31.8	2347	1	US-08-453-703-3
25	1250.5	31.8	2347	2	US-08-456-106-3
26	1250.5	31.8	2347	3	US-08-456-108-3
27	1250.5	31.8	2347	3	US-09-265-577-3
28	1250.5	31.8	2347	4	US-09-633-739-3
29	1249.5	31.7	2319	1	US-08-154-019-1
30	1249.5	31.7	2319	1	US-08-461-333-1
31	1249.5	31.7	2319	3	US-08-464-167-1
32	1249.5	31.7	2319	3	US-09-158-313-1
33	1249.5	31.7	2319	3	US-08-476-798-1
34	1249	31.7	2619	1	US-08-154-019-3
35	1249	31.7	2619	1	US-08-461-333-3
36	1249	31.7	2619	3	US-08-464-167-3
37	1249	31.7	2619	3	US-09-158-313-3
38	1239.5	31.5	4601	1	US-08-470-058-3
39	1239.5	31.5	4601	3	US-09-037-188-3
40	1239.5	31.5	4601	3	US-09-285-310-3
41	1239.5	31.5	4601	3	US-09-285-310-3
42	1220.5	31.0	2259	1	US-08-145-681-5
43	1220.5	31.0	2259	2	US-08-453-703-5
44	1220.5	31.0	2259	2	US-08-456-106-5
45	1220.5	31.0	2259	3	US-08-456-108-5

ALIGNMENTS

RESULT 1  
US-08-520-933-1  
; Sequence 1, Application US/08520933  
; Patent No. 5981194  
; GENERAL INFORMATION:  
; APPLICANT: Jefferies, Wilfred A.  
; APPLICANT: McGeer, Patrick L.  
; APPLICANT: Rothenberger, Sylvia  
; APPLICANT: Food, Michael R.  
; APPLICANT: Yamada, Tatsuo  
; APPLICANT: Kennard, Malcolm  
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins  
; TITLE OF INVENTION: as Diagnostic and Therapeutic Agents  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Bereskin & Parr  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/520,933  
; FILING DATE: August 31, 1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shona S. McDiarmid  
; REGISTRATION NUMBER: 38,798  
; REFERENCE/DOCKET NUMBER: 7685-006  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 416-364-7311



Db 1801 AACACAAACGGCCACAAATCCGAGCCCTGGGCTGTGAGCTCAGGTCAGAGGACTATGAA 1860  
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
Db 1861 CTGCTGTGCCCCAACCGGGCCCCGAGCGAGGTGTCCAGTTTCAGAGCTGCAACCTGGCA 1920  
QY 621 GlnIleProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
Db 1921 CAGATACCAACCCACCGCGTGATGGTCCGCCCGCACCAACATCTTCACCGGTATGGA 1980  
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspHisAsnLysAsnGlyPheLysMet 660  
Db 1981 CTGCTGCAGAGCCCAAGGACCTGTTTGGAGACGACCAATAAGAACGGGTTCAAAATG 2040  
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAla 680  
Db 2041 TTCGACTCTCCAACTATCATGSCCAAGACCTGCTTTCAAGGATGCCCGTCCGGGGC 2100  
QY 681 ValProValGlyGlnLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700  
Db 2101 GTCCCTGTGCGAGAGAAACACCTACCGCGGTGCTGGGGCTGGACTACGTGGCGGG 2160  
QY 701 LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720  
Db 2161 CTGGAAGGGATGTCGTCTAGCAGTGTCTGGGGCGACGCGGCCCGCGGGCGCCCC 2220  
QY 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738  
Db 2221 CTGCTCCCGTGTGTGTCGCCGCCCTCGCGCGCGCTGCTCCCGCGCGGCCCTC 2274

RESULT 2

US-09-285-040-1  
; Sequence 1, Application US/09285040  
; Patent No. 6455494  
; GENERAL INFORMATION:  
; APPLICANT: Jefferies, Wilfred A.  
; APPLICANT: McGeer, Patrick L.  
; APPLICANT: Rothenberger, Sylvia  
; APPLICANT: Food, Michael R.  
; APPLICANT: Yamada, Tatsuo  
; APPLICANT: Kennard, Malcolm  
; TITLE OF INVENTION: Use of p97 and Iron Binding Proteins  
; TITLE OF INVENTION: as Diagnostic and Therapeutic Agents  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bereskin & Parr  
; STREET: 40 King Street West  
; CITY: Toronto  
; STATE: Ontario  
; COUNTRY: Canada  
; ZIP: M5H 3Y2  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/285,040  
; FILING DATE: 01-APR-1999  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gravelle, Micheline  
; REGISTRATION NUMBER: 40,261  
; REFERENCE/DOCKET NUMBER: 7685-032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 416-364-7311  
; TELEFAX: 416-361-1398  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2368 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 61..117  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 118..2274  
US-09-285-040-1  
Alignment Scores:  
Pred. No.: 0 Length: 2368  
Score: 3936.00 Matches: 738  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-10-049-957-4 (1-738) x US-09-285-040-1 (1-2368)  
QY 1 MetArgGlyProSerGlyValAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20  
Db 61 ATGCGGGGTCCGAGCGGGGCTCTGTGGCTGCTCTGGCTCTGGCCACCGTGTCTCGAGGC 120  
QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
Db 121 ATGGAGGTGCGTGTGGTGGCCACCTCGGACCCAGAGCAGCAAGTGCAGCAACATGAGC 180  
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
Db 181 GAGGCGCTTCGGGAAGCGGGCATCCAGCCCTCCCTCTCTGGTCCGGGCGACCTCCGCC 240  
QY 61 AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80  
Db 241 GACCACTGCGTCCAGCTCATCGGGCCAGAGGCTGACGCCATCACTCTGGATGAGGA 300  
QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100  
Db 301 GCCATCTATGAGCGGGAAAGAGCAGCAGCGCTGAGAGCGGTGTGGCGCAAGTGTACGAT 360  
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValArgSerSerHisValThr 120  
Db 361 CAAGAGGTGCGTACCTCTCTATTCGCGGTGTGTGTGTCAGGAGGAGCTCCCATGTGACC 420  
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140  
Db 421 ATTGACCCCTGAAAGGCGTGAAGTCTCCACACGGGCATCAATCGCAGTGGGCTGG 480  
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
Db 481 AACGTGCGCGTGGGTACTCTGGTGGAGAGCGCGCGCTCTCGGTGATGGCTCGCATGTA 540  
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180  
Db 541 CTCAAAAGCTGTCCAGCACTATTTTGGGGCAGCTGCGTCCCGGGGCGAGAGACAGT 600  
QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluClyValCysAsp 200  
Db 601 TACTCTGAGTCCCTCTGTGCGCTCTGACAGGGGTGACAGCTCTGGGGAAGGGGTGTGTGAC 660  
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly 220  
Db 661 AAGAGCCCTCGAGAGATACTACACTACAGCGGGGCTTCCGGTGCCTCGCGGAGGG 720  
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
Db 721 GCAGGGGACGTGGCTTTTGTGAAGCACAGCACGGTACTGGAGAACACCGATGGGAAGACG 780  
QY 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260  
Db 781 CTTCCCTCTCGGGGCGAGGCCCTGTGTTCACAGGACTTCAGCTGTGTGCGCGGATGGT 840  
QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280

```
Db      841  |||||ACCGCGCGCATGTCCACGAGTGGAGGCGAGTCCCATCTGGCCCGGTCGTGCTCAGCC 900
Qy      281  ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGlyGln 300
Db      901  GTGGTGGTCCGGCGCGACACAGATGGGGCTCATCTTCGGGCTGCTCAACGAGGCCAG 960
Qy      301  ArgLeuPheSerHisGluGlySerPheGlnMetPheSerSerGluAlaTyrGlyGln 320
Db      961  CGTCTGTTCCAGCCAGCGAGCGAGCTTCAGATGTTCCAGCTCGAGCCCTATGGCCAG 1020
Qy      321  LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340
Db      1021  AAGGATCTACTCTTCAAGAGCTCTACCTCGAGGCTGTGCCATCGCCACACAGACCTAT 1080
Qy      341  GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360
Db      1081  GAGGCGTGGCTGGGCGCATGAGTACCTGCACGCCATGAAGGGTCTGCTGTGACCCCAAC 1140
Qy      361  ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380
Db      1141  CGGCTGCCCCCTACTCGGCTGGTGTGCTCTCCACTCCCGAGATCCAGAGTGTGGA 1200
Qy      381  AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400
Db      1201  GACATGGCGGTGCTTCGCCCGCGAGCGCTCAAGCCAGAGATCCAGTGGTGTGAGCC 1260
Qy      401  LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420
Db      1261  AAGTCCCCCAACACTGCTGATGAGCGGATCCAGGCTGAGCAGGTGCAGCGTGTGACCTA 1320
Qy      421  SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440
Db      1321  AGTGGCGAGGACATTTACACGCGGGGAGAGTACGGCTGTGTTCCCGCAGCGCGCGAG 1380
Qy      441  HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp 460
Db      1381  CACTATGCCCCGGAAGACAGCAGCAACTCGTACTAGTGGTGGCGCTGGTGGAGCGGAC 1440
Qy      461  SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480
Db      1441  AGCTCCACCGCTTCCACTTGGATGAGTTCGGGGCAAGCGCTCTGCGCAGCGGTTC 1500
Qy      481  GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500
Db      1501  GGCAGCCCTGCAAGCTGGGATGTCCTCGTGGTGGCTTATTCAGAGAGGCTTCATCCGG 1560
Qy      501  ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520
Db      1561  CCCAAGGACTGTGACGTCTCACAGCAGTACGAGGATTTCTCAATGCCAGCTGCTGCC 1620
Qy      521  ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540
Db      1621  GTGAACAAACCCCAAGAACTACCCCTCTCGTGTGTGCACTGTGCGTGGGGACGAGCAG 1680
Qy      541  GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560
Db      1681  GGCCCGCAACAGTGTGTGGGCAACAGCCAGGAGCGGTATTCAGGGTACCGGGCCCTTC 1740
Qy      561  ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580
Db      1741  AGTGCCTGTGTGAGAAATGCGGGTGACGTTCCTTCGTGAGCACAACCGCTCTTGAC 1800
Qy      581  AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600
Db      1801  AACACAAACGGCCACAATTCGAGCCCTGGGCTGTGAGCTCAGGTCCAGAGGACTATGAA 1860
Qy      601  LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
Db      1861  CTGCTGTGCCCCAAGCGGGGCCGAGCCGAGGTGTCCCAATTTGAGCCTGCAACCTGGCA 1920
Qy      621  GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
```

```
Db      1921  CAGATACACACCCACGCGCTGATGTCGGGCCCGACACCAACATCTTCCCGTGTATGGA 1980
Qy      641  LeuLeuAspLysAlaGlnAspLeuPheClyAspAspHisAsnLysAsnGlyPheLysMet 660
Db      1981  CTGCTGGACAAGGCCAGGACCTGTTTGGAGACCAACCAATAAGAACGGGTTCAAAATG 2040
Qy      661  PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680
Db      2041  TTCGACTCTCAACTATCATGGCCAGACCTGTTTCAAGGATGCCACCGTCCGGCGG 2100
Qy      681  ValProValGlyGluLysThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaAla 700
Db      2101  GTGCTGTGCGAGAGAAAACCACTACCGCGCTGGCTGGGCTGGACTAGCTGGCGCGC 2160
Qy      701  LeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGlyAlaPro 720
Db      2161  CTGGAAGGATGTCTCAGCAGTGTCTGGCGCAGCGCCCGCGCGCGCGCGCGCC 2220
Qy      721  LeuLeuProLeuLeuProAlaAlaAlaArgLeuLeuProProAlaLeu 738
Db      2221  CTGCTCCCGCTGCTGCTGCCCGCGCTGCTGCCCGCGCGCGCTC 2274

RESULT 3
5262177-1
;PATENT No. 5262177
;APPLICANT: BROWN J OSEPH P.;ESTIN CHARLES D.;PLOWMAN, GREGORY
;D.;HELLSTROM, KARL E.;ROSE, TIMOTHY M.;HELLSTROM, INGEGRD;
;PURCHIO, ANTHONY F.;HU, SHIU-LOK;PENNATHOR, SRIDHAR
;TITLE OF INVENTION: RECOMBINANT VIRUSES ENCODING THE HHUMAN
;MELANOMA-ASSOCIATED ANTIGEN
;NUMBER OF SEQUENCES: 6
;CURRENT APPLICATION DATA:
;APPLICATION NUMBER: US/07/7230
;FILING DATE: 27-JAN-1987
;PRIOR APPLICATION DATA:
;APPLICATION NUMBER: 827,313
;FILING DATE: 07-FEB-1986
;SEQ ID NO:1:
;LENGTH: 2368
5262177-1

Alignment Scores:
Pred. No.: 0 Length: 2368
Score: 3936.00 Matches: 738
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-049-957-4 (1-738) x 5262177-1 (1-2368)

Qy      1  MetArgGlyProSerGlyAlaLeuTrpLeuLeuLeuAlaLeuArgThrValLeuGlyGly 20
Db      61  ATGCGGGGTCCGAGCGGGCTCTGTGGTGTCTCTCTGCTGCTGCGCACCGTGTCTCGAGGC 120
Qy      21  MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40
Db      121  ATGAGAGTGCCTGCTGCGCCACCTCGGACCCAGAGCAGCAGTGGCGCAACATGAGC 180
Qy      41  GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60
Db      181  GAGCGCTTCCGGGAAGCGGGCATCCAGCCCTCCCTCTCTGCTCGGGCGCACCTCCGCC 240
Qy      61  AspHisCysValGlnLeuLeuAlaAlaGlnAlaAspAlaIleThrLeuAspGlyGly 80
Db      241  GACCACCTGCTCCAGCTCATCGCGCCCGAGAGGTGACGCGCATCATCTCTGATGAGGA 300
Qy      81  AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyValTyrAsp 100
Db      301  GCCATCTATGAGCGGGGAAGAGCAGCGCTGAAGCGGTGGTGGGGAAGTGTACCAT 360
Qy      101  GlnGluValGlyThrSerTyrTyrAlaValAlaValArgSerSerHisValThr 120
```



Db 361 CAAGAGGTGGTACCTCCTATTACCGCGTGGCTGTGTCAGGAGGAGCTCCCATGTGACC 420  
Qy 121 IleAspThrLeuIysGlyValIysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140  
Db 421 ATTGACACCTGAAGCGGTGAAGTCTCTCCACACGCGCATCAATCGCACAGTGGGCTGG 480  
Qy 141 AnnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
Db 481 AACGTGCCGTGGGTACCTGTGGGAGAGCGCGCTCTCGGTGATGGGCTCGCATGTA 540  
Qy 161 LeuIysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180  
Db 541 CTCAAAGCTGTGACGACATATTTTGGGGGAGCTGGTCCCGGGGAGAGACCACT 600  
Qy 181 TyrSerGluSerLeuCysArgLeuCysArgIysAspSerSerGlyGlyValCysAsp 200  
Db 601 TACTCTGAGTCCCTCTGTGCGCTCTGACGGGTGACAGCTCTGGGGAGGGGTGTGAC 660  
Qy 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly 220  
Db 661 AAGAGCCCCCTGGAGAGATACAGACTACAGCGGGGCTTCGGGTGCTGGCGGAAGG 720  
Qy 221 AlaGlyAspValAlaPheValIysHisSerThrValLeuGluAsnThrAspGlyIysThr 240  
Db 721 GCAGGGGAGCTGGCTTTTGTGAAGCACAGACAGCTGAGTGGAGACACGGATGGAGAGC 780  
Qy 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260  
Db 781 CTTCCCTCTGGGGCAGCGCCCTGTGTGTCAGAGACTTCGAGCTGTGTGCGGGATGT 840  
Qy 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280  
Db 841 AGCGGGGCGATGTCAACGAGTGGAGGAGTGCATCTGCGCGGGTGTGCTCACGCC 900  
Qy 281 ValValValArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300  
Db 901 GTGGTGTTCGGGCCACACAGATGGGGCTCATCTTCGGTGTCTCAACGAGGCCAG 960  
Qy 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320  
Db 961 CGTCTGTTCAGCCACGAGGGCAGCAGCTTCAGATGTTCCAGCTCTGAGGCTATGGCCAG 1020  
Qy 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340  
Db 1021 AAGGATCTACTTCAAGACTCTACTCGGAGCTTGTGCCATCGCCACACAGACCTAT 1080  
Qy 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360  
Db 1081 GAGGCGTGGTGGGCCCATGAGTACCTGCACCGCATGAAGGGTCTGCTCTGTGACCCCAAC 1140  
Qy 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380  
Db 1141 CGGTCGCCCTTACCTGCGCTGTGTGTCTCTCCACTCCGAGATCCAGAAGTGTGA 1200  
Qy 381 AspMetAlaValAlaPheArgGlnArgLeuLysProGluIleGlnCysValSerAla 400  
Db 1201 GACATGCGCGTGGCTTCGCGCGCAGCGCTCAAGCCAGAGATCCAGTGGCTGTGAGCC 1260  
Qy 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeu 420  
Db 1261 AAGTCCCCCAACACTGCATGGAGCGGATCCAGGCTGAGAGTGCAGCTGTGACCTTA 1320  
Qy 421 SerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440  
Db 1321 AGTGGCGAGACATTTACAGCGGGGAGAGAGTACGGCTGGTCCCGCAGCGCGAG 1380  
Qy 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgAsp 460  
Db 1381 CACTATGCCCGGAGACAGACAACTCGTACTAGTGGTGGCCGTGGTGGAGCGGAC 1440  
Qy 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
Db 1441 AGCTCCACGCTTACCTTGGATGAGCTTCGGGGCAAGCGCTCTCTGCCACCGCGTTTC 1500

Qy 481 GlySerProAlaGlyTyrAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
Db 1501 GGCAGCCCTGCAGGCTGGGATGTCCCGTGGGTGCCCTATTTCAGAGAGCTTCATCCGG 1560  
Qy 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
Db 1561 CCCAAGGACTGTGACGTCTCACAGCAGTACGAGTATCTTCAATGCCAGCTGCGTGGCC 1620  
Qy 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
Db 1621 GTGAACAACCCCAAGAACTACCCCTCTCGCTGTGTGCACTGTGCTGGGGACGAGCAG 1680  
Qy 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrArgGlyAlaPhe 560  
Db 1681 GSCCGCAACAGTGTGTGGCAACAGCCAGGAGCGGTATTACGGCTACCGCGCGCTTC 1740  
Qy 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580  
Db 1741 AGGTGCTGTGTGGAATGCGGGTGACGTTCCTTCGTGAGGCACACACCCGTCTTTGAC 1800  
Qy 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600  
Db 1801 AACACAAAGCGCACAAATCCAGGCCCTGGGTGTGTGAGCTCAGGTACAGGACTATGAA 1860  
Qy 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
Db 1861 CTGCTGTGCCCCCAACGGGGCCGAGCGAGGTGTCCAGTTTGCAGCTTGCACACCTGGCA 1920  
Qy 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
Db 1921 CAGATACCAACCCCGCGCTGATGTCCGGCCCGACACCAATCTTCAACCTGTATGA 1980  
Qy 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660  
Db 1981 CTGCTGGACAGGCCCGCAGGACTGTTTGGAGACGACCAACATAGAACGGGTTCAAAATG 2040  
Qy 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680  
Db 2041 TTGAGCTCTCTCAACTATCATGCGCAAGACCTGTCTTCAAGGATGCCACCGTCCGGCG 2100  
Qy 681 ValProValGlyLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700  
Db 2101 GTGCTGTGGAGAGAAACCACTACCGCGCTGGCTGGGCTGGACTACGTGGCGCG 2160  
Qy 701 LeuGluGlyMetSerSerGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720  
Db 2161 CTGGAGGGATGTCTCTCAGCAGTGTCTGGGCGCAGCGGCCCGCGCGCGCGCGCC 2220  
Qy 721 LeuLeuProLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738  
Db 2221 CTGCTCCGCTGTCTGTCCCGCCCTCGCGCGCGCTGTCTCCCGCGCGCGCTC 2274

## RESULT 4

US-08-175-158A-1

; Sequence 1, Application US/08175158A

; Patent No. 5986067

; GENERAL INFORMATION:

; APPLICANT: FUNK, Walter D.

; APPLICANT: MAGILLIVRAY, Ross T.A.

; APPLICANT: MASON, Anne B.

; APPLICANT: WOODWORTH, Robert C.

; TITLE OF INVENTION: RECOMBINANT TRANSFERRINS, TRANSFERRIN HALF-

; TITLE OF INVENTION: MOLECULES AND MUTANTS THEREOF

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE &amp; COCKFIELD

; STREET: 60 State Street, suite 510

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII text  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US 07/832,029  
 FILING DATE: 06-FEB-1992  
 FILING DATE: 28-DEC-1993  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/832,029  
 FILING DATE: 06-FEB-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: DeConti, Giulio A.  
 REGISTRATION NUMBER: 31,503  
 REFERENCE/DOCKET NUMBER: UVI-005CP2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2327 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 31..2124  
 FEATURE:  
 NAME/KEY: mat\_peptide  
 LOCATION: 88..2124  
 US-08-175-158A-1

Alignment Scores:  
 Pred. No.: 1,64e-125 Length: 2327  
 Score: 1270.50 Matches: 313  
 Percent Similarity: 55.29% Conservative: 121  
 Best Local Similarity: 39.87% Mismatches: 225  
 Query Match: 32.28% Indels: 126  
 DB: 2 Gaps: 34

US-10-049-957-4 (1-738) x US-08-175-158A-1 (1-2327)

QY	1	MetArgGlyProSerGlyAlaLeuTrpLeu-----LeuLeuAlaLeuArgThrValLeu	18
DB	31	ATGAGGCTCGCGTGGAGCCCTCTGGTCTCGCGCGCTCTGGGCTGTGTCTGGCTGTC	90
QY	19	GlyGlyMetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsn	38
DB	91	CCTGATAAAACTGTGAGATGGTGTGCAGTGTCCGAGCATGAGGCCACTAAGTGC-----	144
QY	39	MetSerGluAlaPheArgGlu-----AlaGlyLeGlnProSerLeu	52
DB	145	-----CAGAGTTCCCGCACCACCATATGAAAGCGTCATTCCATCCGATGGTCCCGAGTGT	198
QY	53	LeuCysValArgGlyThrSerAlaAspHisCysValGlnLeuLeuAlaAlaGlnGluAla	72
DB	199	GCTTGTGTGAAGAACCTCTTACCTTGTATTCATTCAGGGCCATTTCGGCAACGAAGCG	258
QY	73	AspAlaIleThrLeuAspGlyGlyAlaIleTyrGluAlaGly-----LysGluHisGlyLeu	91
DB	259	GATGCTGTGACACTGATGATGAGTGTGGTCTATGATGCTTACTTGGCTCCCAATAACCTG	318
QY	92	LysProValValGlyGluValTyr-----AspGlnGluValGlyThrSerTyrTrpAla	109
DB	319	AAGCCCTGTGGTGGCAGAGTTCATGGGTCAAAAGAGGATCCACAGACTTTCTATTATGCT	378
QY	110	ValAlaValValArgSerSerHisValThrIleAspThrLeuLysGlyValLysSer	129
DB	379	GTTCCTGTGTGAAGAGGATATGTGCTTCAGATGAACACGCTTCGAGGCAAGAGTCC	438
QY	130	CysHisThrGlyIleAsnArgThrValGlyTyrAsnValProValGlyTyrLeuValGlu	149

QY 481 GlySerProAlaGlyTyrAspValProValGlyAlaLeuLeuGlnArgGlyPheIleArg 500  
Db 1450 GGCAGAACCGCTGGTGAACATCCCATGGCTGCTC-----TACAATAAG 1497  
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
Db 1498 ATCAACCATGC-----AGATTGATGAATTTTTTCAGTGAAGTTTGCCCT 1545  
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
Db 1546 GGGTCTAAGAAA-----GACTCCAGTCTCTGTAAGCTGTATGGC-----TCA 1590  
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrGlyTyrArgGlyAlaPhe 560  
Db 1591 GGCCTAAACCTGTGTGAACCAACAAAGAGGATACCTACCGCTACACAGCGCTTC 1650  
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrValPheAsp 580  
Db 1651 AGGTGTCTGGTTGAGAAG---GGAGATGTGGCTTTTGTGAACACACAGACTGTCACAG 1707  
QY 581 AsnThrAsnGlyHisAsnSerGluProTyrAlaAlaGluLeuArgSerGluAspTyrGlu 600  
Db 1708 AACACTGGGGGAAAAAACCTGTATCATCGCTGAAGTGAAGTGAAGTGAAGTGAAG 1767  
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
Db 1768 TTGCTGTGCTTGTATGGTACAGAAACCTGTGGAGGAGTATCGGAACCTGCCACCTGGCC 1827  
QY 621 GlnIleProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
Db 1828 AGACGCCCAATCACCTGTGGTGCACACGGAAGATGAAGAACTGCTCCACAAAG 1884  
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsn-----LysAsnGlyPhe 658  
Db 1885 ATATTACCTCAACAGACGACCTATTTTGAAGACGTAACCTGCTGCTGGCAACTTT 1944  
QY 659 LysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrVal 678  
Db 1945 TGTTGTTCGGTGGCAACCC-----AAGGACCTCTCTGTTACAGATGACACAGTA 1995  
QY 679 ArgAlaValProValGlyGluLysThrTyrArgGlyTyrLeuGlyLeuAspTyrVal 698  
Db 1996 TGTTGGCCCAACTTCATCAGACAAACATATGAAAAATACTTAGGAGAAGATATGTC 2055  
QY 699 AlaAlaLeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGly 718  
Db 2056 AGGCTGTGTGTAACTG-----AGAAATGCTCCACCTCATCA----- 2094  
QY 719 AlaProLeuProLeuLeuLeuProAlaLeuAlaAlaArgLeu----- 733  
Db 2095 -----CTCCTGGAGGCTGCACCTTCGCTAGACCTTAAATCTCAGAGT 2139  
QY 734 ----LeuProPro 736  
Db 2140 AGGGCTGCCACCA 2152

## RESULT 5

US-08-848-760B-25  
; Sequence 25, Application US/08848760B  
; Patent No. 6248721  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Lung-Ji  
; TITLE OF INVENTION: Animal Model For Evaluation Of Vaccines  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: Florida  
; COUNTRY: United States of America  
; ZIP: 32606  
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/848,760B  
APPLICATION NUMBER: 08/838,702  
CLASSIFICATION: <Unknown>  
FILING DATE: 09-APR-1997  
APPLICATION NUMBER: 08/838,702  
ATTORNEY/AGENT INFORMATION:  
NAME: PACE, DORAN R.  
REGISTRATION NUMBER: 38,261  
REFERENCE/DOCKET NUMBER: CNG-100C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (352) 375-8100  
TELEFAX: (352) 372-5800  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2097 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "DNA"  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-08-848-760B-25  
Alignment Scores:  
Pred. No.: 1,778-125 Length: 2097  
Score: 1269.50 Matches: 304  
Percent Similarity: 56.12% Conservative: 118  
Best Local Similarity: 40.43% Mismatches: 221  
Query Match: 32.25% Indels: 109  
DB: 3 Gaps: 32  
US-10-049-957-4 (1-738) x US-08-848-760B-25 (1-2097)  
QY 1 MetArgGlyProSerGlyAlaLeuThrLeu-----LeuLeuAlaLeuArgThrValLeu 18  
Db 1 ATGAGCTCGCCGCTGGAGCCCTGTGGTCTCGCCGCTCTGGGCTGTGCTGGCTGTC 60  
QY 19 GlyGlyMetGluValArgTyrCysAlaThrSerAspProGluGlnHisLysCysGlyAsn 38  
Db 61 CTGTATAAACTGTGAGTGGTGTGCAGCTGTGGAGCATGAGGCCACTAAGTGC----- 114  
QY 39 MetSerGluAlaPheArgGlu-----AlaGlyIleGlnProSerLeu 52  
Db 115 -----CAGAGTTTCCGGGACCATATGAAAGCGTCATTCCCATCGATGGTCCCAGTGT 168  
QY 53 LeuCysValArgGlyThrSerAlaAspHisCysValGlnLeuLeuAlaGlnGluAla 72  
Db 169 GCTTGTGTGAAGAAAGCCCTCTACCTTGATTGCATCAGGGCCATTGCGGCAACGAGCG 228  
QY 73 AspAlaIleThrLeuAspGlyGlyAlaIleTyrGluAlaGly---LysGluHisGlyLeu 91  
Db 229 GATGCTGTGACACTGGATGCAGTTTGGTGTATGATGCTTACTTGGCTCCCAATAACCTG 288  
QY 92 LysProValValGlyGluValTyr-----AspGlnGluValGlyThrSerTyrTyrAla 109  
Db 289 AAGCCTGTGGTGGCAGAGTTCTATGGTCAAAAGAGGATCCACAGACTTCTATTATGCT 348  
QY 110 ValAlaValValArgSerSerHisValThrIleAspThrLeuLysGlyValLysSer 129  
Db 349 GTTGTGTGGTGAAGAGGATAGTGGCTTCCAGATGAACACCATTCGAGGCAAGAGTCC 408  
QY 130 CysHisThrGlyIleAsnArgThrValGlyTyrAsnValProValGlyTyrLeuValGlu 149  
Db 409 TGCCACAGGGTCTAGGCGAGGTCCGCTGGGTGGTGAACATCCCATAGGCTTACTTTAC--- 465  
QY 150 SerGlyArgLeuSerValMetGlyCysAsp-----ValLeuLys 162

466 -----TGTGACTTACCTGAGCCACGTAACCTCTTGAGAAA 501  
163 AlaValSerAspTyrPheGlySerCysValProGlyAlaGlyGluThrSerTyrSer 182  
502 CGAGTGGCAATTTCTTCGCGAGCTGTGCCCTTGTGCGAGTGGAGCGACTTCCCC 561  
183 GluSerLeuValCysArgGlyAspSerSerGlyGluGlyValCysAspLysSer 202  
562 CAG---CTGTGCAACTGTGTCCAGG-----TGTGGCTGTCTCC 597  
203 ProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGly 222  
598 ACCCTTAACCAATACTTCGGCTACTCGGAGCCTTCAAGTGTCTGAAGATGTCTGG 657  
223 AspValAlaPheValIleHisSerThrValLeuGluAenThrAspGlyLysThrLeuPro 242  
658 GATGTGGCTTTGTCAAGCACTCGACTATATTGTGAGAACTTGGCAACAAGGCT----- 711  
243 SerTyrGlyGlnAlaLeuLeuSerGlnAsp---PheGluLeuLeuCysArgAspGlySer 261  
712 -----CACAGGACCAAGTATGAGCTCTTGTCTAGCAACACC 750  
262 ArgAlaAspValThrGluTyrArgGlnCysHisLeuAlaArgValProAlaHisAlaVal 281  
751 CGGAAGCCGGTAGATGAATACAAAGACTGCACTTGGCCAGGTCCCTTCTCATACCGTC 810  
282 ValValArgAlaAspThrAspGlyGly-----LeuIlePheArgLeuLeuAsnGlu 298  
811 GTGGCCCGCA-----AGTATGGCGCGCAAGGAGGACTTGTATCTGGGAGCTTCTCAACCA 864  
299 GlyGlnArgLeuPheSerHisGluGlySer---SerPheGlnMetPheSerSerGluAla 317  
865 GCCCAGGAAACATTTTGGCAAGCAAAATCAAAAGAAATTCCAACTATTCAAGCTCT---CCT 921  
318 TyrGlyGlnLysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThr 337  
922 CATGG---AAGGACCTGCTTTTAAGGACTCTGCCACCGGGTTTTTAAAGTCCCCCA 978  
338 Gln---ThrTyrGluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeu 356  
979 AGGATGGATGCCAAGATGTACCTGGCTGTAGTATGTCACTGCCATCCGGAATCTACGG 1038  
357 -----CysAspProAsnArgLeuProTyr 365  
1039 GAAGGCACATGCCAGACCCCAACAGATGAATCAAGCT----- 1080  
366 LeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGlyAspMetAlaValAla 385  
1081 GTGAAGTGTGTGCGCTGAGCCACCAAGAGAGGCTCAAGTGTGTGATGAGTGGAGTGT--- 1137  
386 PheArgArgGlnArgLeuLysProGluIleGlnCysValSerAlaLysSerProGlnHis 405  
1138 -----AACAGTGTAGGGAATAAGTAGTGTGTATCAAGCAGACACCCGAAGAC 1185  
406 CysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeuSerGlyGluAspIle 425  
1186 TCATCGCCCAAGATCATGATGAGAGAGCTGTATGCCATGAGCTTGTGATGGAGGTTGTC 1245  
426 TyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGluHisTyrAlaProGlu 445  
1246 TACATAGCGGGC---AAGTGTGCTGTGGTGCCTGTCTTGGCAGAAAACCTACAATAAGAGC 1302  
446 AspSerSerAsn-----SerTyrTyrValValAlaValAlaValArgArgAsp 460  
1303 GATAATGTGAGGATACACAGGACGGGTATTTTGTGTAGCAGTGTGTGAAGAATCA 1362  
461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480  
1363 GCTTTCTGAC---CTACCTGGGACCAATCTGAAGGCGAAGAGTCTCTCCATACGCGAGTT 1419  
481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500  
1420 GGCAGAACCGCTGGCTGGAAACATCCCATGGGCTGCTC-----TACAATAAG 1467

QY 501 ProlysAspCysAspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValPro 520  
Db 1468 ATCAACCACTGC-----AGATTTGATGAATTTTTCAGTGAAGGTGTGCCCTT 1515  
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540  
Db 1516 GGGTCTAAGAAA-----GACTCCAGTCTCTGTAAAGCTGTGTATGGGC-----TCA 1560  
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560  
Db 1561 GGCCTAAACCTGTGTGAACCCACACAAAGAGGAGTACTACGGCTACACAGGCGCTTTC 1620  
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrValPheAsp 580  
Db 1621 AGGTGTCTGGTTTGAAG---GGAGATGTGGCTTTGTGAACACACAGACTGTCCACAG 1677  
QY 581 AsnThrAsnGlyHisAsnSerGluProTyrAlaAlaGluLeuArgSerGluAspTyrGlu 600  
Db 1678 AACACTGGGGGAAAAACCTGTATCCATGGCTAAGATCTGAATGAAAAAGACTATGAG 1737  
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620  
Db 1738 TTGCTGTCCCTTGTATGATGATCCAGGAAACCTGTGGAGGAGTATGCGAACTGCCACCTGGCC 1797  
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640  
Db 1798 AGAGCCCGAATACGCTGTGTGTACACGGAAGATGAAGAAAGCT---TGGCTCCACAG 1854  
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsn-----LysAsnGlyPhe 658  
Db 1855 ATATTACCTCAACAGCAGCACCATTATTTGGAAGCAACGTAACCTGCTCGGGCAACTTT 1914  
QY 659 LysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrVal 678  
Db 1915 TGTGTGTTCGGTCGGAACCC-----AAGGACCTTCTGTGTGAGAGATGACACAGTA 1965  
QY 679 ArgAlaValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrVal 698  
Db 1966 TGTGTGGCAACCTTCATGACAGAAACACATATGAAAATACTTAGGAGAAAGATATGTC 2025  
QY 699 AlaAlaLeuGluGlyMetSerSerGlnGlnCysSer 710  
Db 2026 AAGGCTGTGTGTAACCTG-----AGAAAATGCTCC 2055

## RESULT 6

US-08-145-681-1  
; Sequence 1, Application US/08145681  
; Patent No. 5571691

## GENERAL INFORMATION:

; APPLICANT: Conneely, Orla M.  
; APPLICANT: Headon, Denis R.  
; APPLICANT: O'Malley, Bert W.  
; APPLICANT: May, Gregory S.  
; TITLE OF INVENTION: Production of Recombinant Lactoferrin  
; TITLE OF INVENTION: and Lactoferrin Polypeptides Using cDNA Sequences in  
; TITLE OF INVENTION: Various Organisms

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Baker & Botts, L.L.P.

; STREET: 910 Louisiana St

; CITY: Houston

; STATE: Texas

; COUNTRY: USA

; ZIP: 77002

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/145,681

; FILING DATE:

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcgregor, Martin L.
; REGISTRATION NUMBER: 29,329
; REFERENCE/DOCKET NUMBER: 19928-0125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713/229/1874
; TELEFAX: 713/229/1522
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: H. sapiens
; US-08-145-681-1

```

## Alignment Scores:

```

Pred. No.: 1,07e-124 Length: 2360
Score: 1263.00 Matches: 306
Percent Similarity: 52.23% Conservative: 116
Best Local Similarity: 37.87% Mismatches: 245
Query Match: 32.09% Indels: 142
DB: 29 Gaps: 29

```

US-10-049-957-4 (1-738) x US-08-145-681-1 (1-2360)

```

QY 3 GlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGlyMetGlu 22
DB 52 GGGCCCTCGGACTGTGTCTGGCTG-----GCCGTAGAGA-AGGACT 92
QY 23 ValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSerGluAla 42
DB 93 GTTCAGTGTGTGACCTATCCCAACCCGAGGCCACAAATGCTTCCAATGGCAAGGAAT 152
QY 43 PheArgGluAlaGlyLeuGlnProSerLeuLeuCysValArgGlyThrSerAlaAspHis 62
DB 153 ATGAGAAGAGTGGGT---GGCCCTCTCTGAGCTGCATTAAGAGAGACTCCCCATCCAG 209
QY 63 CysValGlnLeuLeuAlaGlnGluAlaAspAlaIleThrLeuAspGlyGlyAlaIle 82
DB 210 TGTATCCAGCCATTCGGGAACACAGGCCGATGCTGTGACCTTCATGTGTTCATA 269
QY 83 TyrGluAlaGly---LysGluHisGlyLeuLysProValValGlyGluValTyr----- 99
DB 270 TAGAGGCGAGCGCTGGCCCTTACAACTCGACCTGTAGCGCGGAGTCTACGGGACC 329
QY 100 AspGlnGluValGlyThrSerTyrTyrAlaValAlaValValArgSerSerHisVal 119
DB 330 GAAAGACAGCCAGCAACTCATATTATGCGTGTGTGTGAAGAGGGCGGCGAGCTTT 389
QY 120 ThrIleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGly 139
DB 390 CAGCTGAACGAAGTCAAGGTCTGAAGTCTGCGCACACAGGCTTCGCGAGGCGGTGA 449
QY 140 TrpAsnValProValGlyTyrLeu----- 147
DB 450 TGAATGTGCTATAGGACACTTCGTCATTCTTGAATTGGACGGGTCCACCTGAGCCC 509
QY 148 ValGluSerGlyArgLeuSerValMetGlyCysAspValLeuLysAlaValSerAspTyr 167
DB 510 ATTAGGCCA-----GCTGTGCGCAGGTTC 533
QY 168 PheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSerGluSerLeuCysArg 187
DB 534 TTCTAGCCAGCTGTGTTCCTCGGTGCATTAAGGACAGTTC---CCCAACCTGTGTGCG 590
QY 188 LeuCysArgGlyAspSerGlyGluGlyValCysAspLysSerProLeuGluArgTyr 207
DB 199 TCTAATCTCTGCTCTGTGTATTGGCGACGAGCGGTGAGAAATAGTCGCTGCC 1634

```

```

DB 591 CTGTGTGCGGGG-----ACAGGGGAAACAAATGTGCTCTCTCTCCAGAACCGTAC 644
QY 208 TyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGlyAspValAlaPheVal 227
DB 645 TTCAGCTACTCTGTGCTCTCAAGTCTCTGAGAGACGGGGCTGGACAGCTGGCTTTATC 704
QY 228 LysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuProSerTrpGlyGlnAla 247
DB 705 AGAGAGAGCACAGTGTCTGAGGACCTGTCTGAGAGCT----- 743
QY 248 LeuLeuSerGlnAspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGlu 267
DB 744 ---GAAAGGACGAGTATGAGTTACTCTGCCAGACACACATCGGAAGCAGTGGACAAG 800
QY 268 TrpArgGlnCysHisLeuAlaArgValProAlaHisAlaValValAlaArgAla---Asp 286
DB 801 TTCAAAGACTGCCATCTGCCCGGGTCCCTTCTCATGCCGTGTGTGCACGAAGTGTGAAT 860
QY 287 ThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGlnArgLeuPheSerHisGlu 306
DB 861 GGCAAGGAGGATGCCATCTGGAATCTTCTCCGCGGACACAGGAAAGTTTGGAAAGGAC 920
QY 307 GlySer---SerPheGlnMetPheSerSerGluAlaTyrGlyGlnLysAspLeuPhe 325
DB 921 AAGTCACCGAAATTCAGCTCTTGGCTCCCTAGT---GGGCGAGAAAGATCTGCTGTTC 977
QY 326 LysAspSerThr-----SerGluLeuValPro---IleAlaThrGlnThrTyrGlu 341
DB 978 AAGGACTCTGCCATTTGGGTTTTCGAGGTTGCCCGAGGATAGATCTCTGGGCTGTAC--- 1034
QY 342 AlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsnArg 361
DB 1035 ---CTTGGCTCCGGCTACTTCTGCTCATCCAGAACTTCAGGAAAAGTGAAGGAGAA 1088
QY 362 LeuPro-----ProTyrLeuArgTrpCysValLeuSerThrProGluLeuGlnLys 378
DB 1089 GTGGCTCCCGCGCTGCGCGGTCTGTGTGTGCGGTGGCGGAGCAGGAGCTGCCCAAG 1148
QY 379 CysGlyAspMetAlaValAlaPheArgArgGlnArgLeuLysProGluLeuGlnCysVal 398
DB 1149 TGTACCAGTGGAGT-----GGCTTGAGCAAGGCGGTGACCTGCTCC 1193
QY 399 SerAlaLysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaVal 418
DB 1194 TCGGCTCCACACAGAGACTGCTGCGCTGGTCTGAAAGGAGAGCTGATGCCATG 1253
QY 419 ThrLeuSerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAla 438
DB 1254 AGTTTGGATGGAGGATATGTGTACACTGCAGGC---AAATGTGTTTGTGCTGTCTG 1310
QY 439 GlyGluHisTyrAlaProGluAspSerSerAsn----- 449
DB 1311 GCAGAGAACTCAAAATCCCAAGAGTGCCTGATCTTAATCTGTGTGATAGACCT 1370
QY 450 ---SerTyrTyrValValAlaValValArgAspSerSerHisAlaPheThrLeu 467
DB 1371 GTGGAAGGATATCTTGTCTGCGCGTGGTGTAGAGA---TCAGACACTAGCTTACTCG 1427
QY 468 AspGluLeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAlaGlyTrpAsp 487
DB 1428 AACTCTGTGAAGGCAAGAAAGTCTTGCACACCGCGGTGGACAGGACTCGAGCTGGAAT 1487
QY 488 ValProValGlyAlaLeuIleGlnArgGlyPheIleArgProLysAspCysAspValLeu 507
DB 1488 ATCCCATCTGGGCTGCTC-----TTCAACACGAGCGGTCTCTGC----- 1526
QY 508 ThrAlaValSerPhePheAsnAlaSerCysValProValAsnAsnProLysAsnTyr 527
DB 1527 ---AAATTTGATGATATTTAGTCAAGCTGTGCTTGGCTGTGACCCCGAGA----- 1577
QY 528 ProSerSerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGly 547
DB 1578 ---TCTAATCTCTGCTCTGTGTATTGGCGACGAGCGGTGAGAAATAGTCGCTGCC 1634

```









QY 450 -----SerTyrTyrValValAlaValAlaValArgArgAspSerSerHisAlaPheThrLeu 467  
Db 1371 GTGGAAGGATATCTGCTGGCGGTGGTTAGGAGA---TCAGACACTACGCTTACCTGG 1427  
QY 468 AspGluLeuArgGlyArgSerCysHisAlaGlyPheGlySerProAlaGlyTrpAsp 487  
Db 1428 AACTCTGTGAAGAGAGAGTCTGCCACCGCGTGGACAGGACTGCGAGCTGGAAT 1487  
QY 488 ValProValGlyAlaLeuLeuGlnArgGlyPheAlaArgProLysAspCysAspValLeu 507  
Db 1488 ATCCCATGGCTGCTC-----TTCAACACGAGCGGCTCCTGC----- 1526  
QY 508 ThrAlaValSerGluPhePheAenAlaSerCysValProValProValAsnAsnProLysAsnTyr 527  
Db 1527 ---AAATTTGATGAATATTTCAGTCAAGAGTGTGCCCTGGGTCTGACCCGAGA----- 1577  
QY 528 ProSerSerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGly 547  
Db 1578 ---TCTAACTCTGTGCTGTGTATTGGCGAGCAGCAGGAGTGAATAGTGGTGGCC 1634  
QY 548 AsnSerGlnGluArgTyrTyrGlyTyrArgGlyValaPheArgCysLeuValGluAenAla 567  
Db 1635 AACAGCAATGAGAGATACTACGGCTACACTGGGGCTTTCCGGTGGCTGGCTGAGATGCT 1694  
QY 568 GlyAspValAlaPheValArgHisThrValPheAspAsnThrAsnGlyHisAsnSer 587  
Db 1695 GGAGAGCTTGCATTTGTGAAGATGTCACTGTCTTGAGAACACTGATGGAATAACAAT 1754  
QY 588 GluProTTPAlaAlaGluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAla 607  
Db 1755 GAGGCATGGCTGAAGATTTGAAGCTGGCAGACTTTGGCGTGTGTGCTGAGTGCAGAA 1814  
QY 608 ArgAlaGluValSerGlnPheAlaAlaCysAsnAlaGlnIleProProHisAlaVal 627  
Db 1815 CGGAACCTGTGACTGAGGTAGAGCTGCATCTTGCATGCCCCCGAATCATGCCGIG 1874  
QY 628 MetValArgProAspThrAsnIlePheThrValTyrGlyLeuLeuAspLysAlaGlnAsp 647  
Db 1875 GTGTCTCGGATGAT---AAGGTGGAACGCTGAAACAGGTGTCTGCCACCAACAGGCT 1931  
QY 648 LeuPheGlyAspAspHisAsnLysAsnGly-----PheLysMetPhe 661  
Db 1932 AAATTTGGG-----AGAAATGGATCTGACTGCCCGGACAAAGTTTGTCTTATTC 1979  
QY 662 AspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAlaVal 681  
Db 1980 CAGTCTGAAC-----AAAAACCTTCTGTTCAATGACAACTGAGTGTCTGGCC 2030  
QY 682 ProValGlyGluLysThrThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaAlaLeu 701  
Db 2031 AGACTCCATGGCAAAACACATATGAAAAATATTGGGACCACAGTATGTCCGAGGCATT 2090  
QY 702 GluGlyMet----- 704  
Db 2091 ACTAATCTGAAAAAGTGTCAACCTCCCTCCCTGGAAGCCTGTGAATTCCTCAGGAAG 2150  
QY 704 ----- 704  
Db 2151 TAAACCGAAGAGATGGCCAGCTCCCAAGAAAGCCTCAGGCATTCACTGCCCCAGC 2210  
QY 705 ---SerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGly-----AlaProLeu 721  
Db 2211 TCTTCTCCCGAGGTGTGTGGGGCTTGGCTCCCTGCTGCTGAAGTGGGATTGCCCATCC 2270  
QY 722 LeuProLeuLeuLeuProAlaLeu 729  
Db 2271 ATCTGCTTACAATTCCCTGTCTGTC 2294

## RESULT 9

US-08-456-106-1

; Sequence 1, Application US/08456106

; Patent No. 5849881

GENERAL INFORMATION:  
APPLICANT: Comeely, Orla M.  
APPLICANT: Headon, Denis R.  
APPLICANT: O'Malley, Bert W.  
APPLICANT: May, Gregory S.  
TITLE OF INVENTION: Production of Recombinant Lactoferrin  
TITLE OF INVENTION: and Lactoferrin Polypeptides Using cDNA Sequences in  
NUMBER OF INVENTIONS: Various Organisms  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/456,106  
FILING DATE: Concurrently herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/145,681  
FILING DATE: October 28, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Albert P. Halluin  
REGISTRATION NUMBER: 25,227  
REFERENCE/DOCKET NUMBER: 8206-025  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-854-3660  
TELEFAX: 415-854-3694  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2360 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: H. sapiens  
US-08-456-106-1  
Alignment Scores:  
Pred. No.: 1,07e-124 Length: 2360  
Score: 1263.00 Matches: 306  
Percent Similarity: 52.23% Conservative: 116  
Best Local Similarity: 37.87% Mismatches: 245  
Query Match: 32.09% Indels: 142  
DB: 29 Gaps: 29  
US-10-049-957-4 (1-738) x US-08-456-106-1 (1-2360)  
QY 3 GlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGlyMetGlu 22  
Db 52 GGGCCCTCGGACTGTGTCTGGCTG-----GCCGTAGAGA-AGGAGT 92  
QY 23 ValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSerGluAla 42  
Db 93 GTTCAGTGTGTGACCGTATCCCAACCCGAGGCCACAAATGCTTCAATGGCAAGGAAT 152  
QY 43 PheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAlaAspHis 62  
Db 153 ATGAGAAGAGTGCCT---GGCCCTCTCTGTCAGCTGCATAAAGAGAGACTCCCCATCCAG 209  
QY 63 CysValGlnLeuIleAlaAlaGlnGluAlaAspAlaIleThrLeuAspGlyGlyAlaIle 82  
Db 210 TGTATCCAGGCCATTGCGGAAAAACAGGGCCGATGTGTGACCTTGTATGGTGTTCATA 269

QY 83 TyrGluAlaGly---LysGluHisGlyLeuLysProValValGlyGluValTyr----- 99  
DB 270 TACGAGGAGGCTGCGCCCTTACAACTGCGACTGTAGCGCGGAAGTCTACCGGACC 329  
QY 100 AspGlnGluValGlyThrSerTyrTyrAlaValAlaValValArgSerSerHisVal 119  
DB 330 GAAAGACACCGCAACTCCTATTATGCGGTGGCTGGTGAAGAGGGCGGACGCTTT 389  
QY 120 ThrIleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGly 139  
DB 390 CAGCTGAAACGAACTCAAGTCTGAAGTCTGCTCCACACAGGCTTCGACGAGCCGCTGGA 449  
QY 140 TrpAsnValProValGlyTyrLeu----- 147  
DB 450 TGGAAATGCGCTATAGGACACTTCGTCATTCCTGTAATGACCGGTCCACCTGAGCCC 509  
QY 148 ValGluSerGlyArgLeuSerValMetGlyCysAspValLeuLysAlaValSerAspTyr 167  
DB 510 ATTGAGGCA-----GCTGTGGCCAGGTTTC 533  
QY 168 PheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSerGluSerLeuCysArg 187  
DB 534 TTCTCAGCCAGCTGTGTTCCCGGTGCAGATAAAGGACAGTTTC---CCCAACCTGTGTGCG 590  
QY 188 LeuCysArgGlyAspSerSerGlyGluGlyValCysAspLysSerProLeuGluArgTyr 207  
DB 591 CTGTGTGGGG-----ACAGGGAAACAAATGTGCTTCTCTCCAGGACCGTAC 644  
QY 208 TyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGlyAspValAlaPheVal 227  
DB 645 TTCAGCTACTCTGGTGCCTTCAAGTCTGAGAGACGGGCTGGAGAGCTGGCTTTTATC 704  
QY 228 LysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuProSerTyrGlyGlnAla 247  
DB 705 AGAGAGAGCACAGTGTGAGGACCTGTCCAGAGGCT----- 743  
QY 248 LeuLeuSerGlnAspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGlu 267  
DB 744 ---GAAAGGACGATGATGATGTTACTCTGCCAGACACACTCGGAAGCCAGTGGACAAG 800  
QY 268 TrpArgGlnCysHisLeuAlaArgValProAlaHisAlaValValAlaValAla---Asp 286  
DB 801 TTCAAAGACTGCACTCGCCCGGCTCCCTTCTCATGCGGTGTGGCAGCAAGTGTGAAT 860  
QY 287 ThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGlnArgLeuPheSerHisGlu 306  
DB 861 GGCAAGGAGGATGCCATCTGGAATCTTCTCCGCCAGGCACAGAAAGTTTGGAAAGGAC 920  
QY 307 GlySer---SerPheGlnMetPheSerSerGluAlaTyrGlyGlnLysAspLeuPhe 325  
DB 921 AAGTCACCGAATTCAGCTCTTGGCTCCCTAGT---GGCAGAAAGATCTGCTGTTC 977  
QY 326 LysAspSerThr-----SerGluLeuValPro---IleAlaThrGlnThrTyrGlu 341  
DB 978 AAGGACTCTGCCATTGGGTTTTTCGAGGGTGCCTCCGAGGATAGATTCTGGGCTGTAC--- 1034  
QY 342 AlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsnArg 361  
DB 1035 ---CTTGGCTCGGCTACTTCACTGCCATCCAGAACTTGAGGAAAGTGAGGAGGAA 1088  
QY 362 LeuPro-----ProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLys 378  
DB 1089 GTGGCTGCCCGCGGTGCTGTGGTGTGGTGGCGGACGAGGAGCTGGCGAAG 1148  
QY 379 CysGlyAspMetAlaValAlaPheArgGlnArgLeuLysProGluIleGlnCysVal 398  
DB 1149 TGTAAACAGTGGAGT-----GGCTTGAGCAAGGCGGCTGACCTGCTCC 1193  
QY 399 SerAlaLysSerProGlnHisCysMetGluArgIleGlnAlaGluValValAspAlaVal 418  
DB 1194 TCGGCTCCACACAGAGACTGTCATCGCCCTGCTGCTGAAAGGAGGAGTGTATCCCATG 1253

QY 419 ThrLeuSerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAla 438  
DB 1254 AGTTTGGATGAGGATGATGTGTACTGACGCGC---AAATGTGGTTTGGTGCCTGCTG 1310  
QY 439 GlyGluHisTyrAlaProGluAspSerSerAsn----- 449  
DB 1311 GCAGAGAACTACAATCCCAACAAAGCAGTGCACCTGATCTTAAGTGTGTGGATAGACCT 1370  
QY 450 ---SerTyrTyrValValAlaValValArgArgAspSerSerHisAlaPheThrLeu 467  
DB 1371 GTGGAAGGATATCTTGTGTGGCGTGTGTAGAGA---TCAGACACTAGCTTACCTGG 1427  
QY 468 AspGluLeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAlaGlyTyrAsp 487  
DB 1428 AACTCTGTGAAGGACAGAGTCTGCGCACACCGCGGTGGACAGACTGCAGGCTGGAAAT 1487  
QY 488 ValProValGlyAlaLeuIleGlnArgGlyPheIleArgProLysAspCysAspValLeu 507  
DB 1488 ATCCCATGGCGCTGCTC-----TTCAACACAGACGGGCTCCTGCG----- 1526  
QY 508 ThrAlaValSerGluPhePheAsnAlaSerCysValProValAsnAsnProLysAsnTyr 527  
DB 1527 ---AAATTTGATGAATATTTCAGTCAAAGCTGTGCCCTCTGGTCTGACCCGAGA----- 1577  
QY 528 ProSerSerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGly 547  
DB 1578 ---TCTAATCTCTGTGCTGTGTATTGGCCAGCAGAGGGTGAGAAATAGTGCCTGCC 1634  
QY 548 AsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPheArgCysLeuValGluAsnAla 567  
DB 1635 AACAGCAATGAGAGATACTACCGGTACACTGGGGCTTTCCGTGCTGCTGGCTGAGAATGCT 1694  
QY 568 GlyAspValAlaPheValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSer 587  
DB 1695 GGACAGCTGATTTGTGAAGATGTCACTGTCTTGCAGAACACTGATGGAATAACAACT 1754  
QY 588 GluProTrpAlaAlaGluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAla 607  
DB 1755 GAGGCATGGGCTAAGGATTTGAAGTGCACACTTTGGCTGTGCTGTGCTGATGCGCAA 1814  
QY 608 ArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAlaGlnIleProProHisAlaVal 627  
DB 1815 CGGAAGCCTGTGACTGAGGCTAGAAGCTGCCATCTTGCATGCGGCCCGCAATCATGCCGTG 1874  
QY 628 MetValArgProAspThrAsnIlePheThrValTyrGlyLeuLeuAspLysAlaGlnAsp 647  
DB 1875 GTGTCTCGATGGAT---AAGGTGGAAACGCTGAAACAGGTGCTCTCCACCAACAGGCT 1931  
QY 648 LeuPheGlyAspAspHisAsnLysAsnGly-----PheLysMetPhe 661  
DB 1932 AAATTTGGG-----AGAAATGGATCTGACTGCCCGGACAAAGTTTGTCTATTTC 1979  
QY 662 AspSerSerAsnTyrHisGlyAsnLeuPheLysAspAlaThrValArgAlaVal 681  
DB 1980 CAGTCTGAAACC-----AAAAACCTCTGTCTCAATGACCAACACTGAGTGTCTGGCC 2030  
QY 682 ProValGlyGluLysThrThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaLeu 701  
DB 2031 AGACTCCATGCCAAACAAACATATGAAATAATTTGGGACCAACAGTATGTGCGAGGCATT 2090  
QY 702 GluGlyMet----- 704  
DB 2091 ACTAATCTGAAAAAGTGTCAACCTCCCTCTCTGGAAAGCTGTGAATTCCTCAGGAAG 2150  
QY 704 ----- 704  
DB 2151 TAAACCGAAGAGATGGCCCGCAGCTCCCAAGAAAGCCCTCAGCCATTCACTGCCCCAGC 2210  
QY 705 ---SerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGly-----AlaProLeu 721  
DB 2211 TCTTCTCCCGAGGTGTGTGGGGCTTGGCTCTCCCTGCTGAAGGTGGGATTGCCCATCC 2270  
QY 722 LeuProLeuLeuProAlaLeu 729

```

Db      2271 ATCTGCTTACAATTCCTCTGTC 2294
      ::: ||| |||||:::
US-08-456-108-1
; Sequence 1, Application US/08456108
; Patent No. 6100054
; GENERAL INFORMATION:
; APPLICANT: Conneely, Orla M.
; APPLICANT: Heaton, Denis R.
; APPLICANT: O'Malley, Bert W.
; APPLICANT: May, Gregory S.
; TITLE OF INVENTION: Production of Recombinant Lactoferrin
; TITLE OF INVENTION: And Lactoferrin Polypeptides Using cDNA Sequences in
; TITLE OF INVENTION: Various Organisms
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,108
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/145,681
; FILING DATE: October 28, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Albert P. Halluin
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8206-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-3660
; TELEFAX: 415-854-3694
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: H. sapiens
US-08-456-108-1
Alignment Scores:
Pred. No.: 1,07e-124 Length: 2360
Score: 1263.00 Matches: 306
Percent Similarity: 52.23% Conservative: 116
Best Local Similarity: 37.87% Mismatches: 245
Query Match: 32.09% Indels: 142
DB: 3 Gaps: 29
US-10-049-957-4 (1-738) x US-08-456-108-1 (1-2360)
QY      3 GlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGlyMetGlu 22
      |||||:::
Db      52 GGGCCCTCGAGCTGCTCTGGCTG-----GCCGTAGGAGA-AGGAGT 92
      |||||:::
QY      23 ValArgTrpCysAlaThrSerAspProGluGlnHisCysGlyAsnMetSerGluAla 42
      |||||:::
Db      93 GTTCAGTGTGACCGTATCCCAACCCGAGGCCACAAAATGCTTCCAATGGCAAGGAAT 152
      |||||:::
QY      43 PheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAlaAspHis 62
      |||||:::
Db      153 ATGAGAAGAGTGGT---GGCCCTCTCTGAGTCATTAAGAGAGACTCCCCCATCCAG 209
      |||||:::
QY      63 CysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyAlaIle 82
      |||||:::
Db      210 TGTATCCAGGCCATTGCGGAAACAGGCGCGATGCTGTGACCTTGTATGGTGTTCATA 269
      |||||:::
QY      83 TyrGluAlaGly---LysGluHisGlyLeuLysProValValGlyGluValTyr----- 99
      |||||:::
Db      270 TAGGAGGAGGCTGGCCCCCTACAACTGCACCTGTAGCGCGGAAGTCTACGGGACC 329
      |||||:::
QY      100 AspGlnGluValGlyThrSerTyrTyrAlaValAlaValArgSerHisVal 119
      |||||:::
Db      330 GAAAGACACCCAGCAACTCACTATTATGCGTGGTGTGCTGAAGAAGGCGCGACGCTTT 389
      |||||:::
QY      120 ThrIleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGly 139
      |||||:::
Db      390 CAGCTGAACGAACTCGAAGTCTGAAGTCTGCCACAGGCTTCGCAGGACCGCTGGA 449
      |||||:::
QY      140 TrpAsnValProValGlyTyrLeu----- 147
      |||||:::
Db      450 TGGAAATGCTATAGGACACTTCGTCCATTCTTGAATTGGACGGTCCACTGAGCCCC 509
      |||||:::
QY      148 ValGluSerGlyArgLeuSerValMetGlyCysAspValLeuLysAlaValSerAspTyr 167
      |||||:::
Db      510 ATTGAGGCA-----GCTGTGGCCAGGTTT 533
      |||||:::
QY      168 PheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSerGluSerLeuCysArg 187
      |||||:::
Db      534 TTCTCAGCCAGCTGTGTTCCCGTGCAGTAAGGACAGCTTC---CCCAACCTGTGTGCG 590
      |||||:::
QY      188 LeuCysArgGlyAspSerSerGlyGluGlyValCysAspLysSerProLeuGluArgTyr 207
      |||||:::
Db      591 CTGTGTGCGGG-----ACAGGGGAAACAATGTGCTTCTCTCCCGAGNACCGTAC 644
      |||||:::
QY      208 TyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGlyAspValAlaPheVal 227
      |||||:::
Db      645 TTCAGCTACTCTGGTGCCTTCAAGTGTCTGAGAGCGGGGCTGGAGACGTGGCTTTATC 704
      |||||:::
QY      228 LysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuProSerTrpGlyGlnAla 247
      |||||:::
Db      705 AGAGAGAGCACAGTGTGTTGAGGACCTGTCCAGCAGGCT----- 743
      |||||:::
QY      248 LeuLeuSerGlnAspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGlu 267
      |||||:::
Db      744 ---GAAAGGACAGTATGAGTTACTTCCCGCAGACAACACTCGGAGCCAGTGGACAAG 800
      |||||:::
QY      268 TrpArgGlnCysHisLeuAlaArgValProAlaHisAlaValValArgAla---Asp 286
      |||||:::
Db      801 TTCAAAGACTGCCATCTGCGCGGGTCCCTTCTCATGCGTGTGTGGCACGAAGTGTGAAT 860
      |||||:::
QY      287 ThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGlnArgLeuPheSerHisGlu 306
      |||||:::
Db      861 GGCAAGGAGGATGCCATCTTCTCCCGCAGGACACAGAAAAGTTTGGAAAGGAC 920
      |||||:::
QY      307 GlySer---SerPheGlnMetPheSerSerGluAlaTyrGlyGlnLysAspLeuLeuPhe 325
      |||||:::
Db      921 AAGTACCCGAAATTCAGCTCTTTGGTCCCTCCCTAGT---GGCAGAAAGATCTGCTGTTTC 977
      |||||:::
QY      326 LysAspSerThr-----SerGluLeuValPro---IleAlaThrGlnThrTyrGlu 341
      |||||:::
Db      978 AAGGACTCTGCCATTGGGTTTTCGAGGGTGGCCCCGAGGATAGATTCTGGGCTGTAC--- 1034
      |||||:::
QY      342 AlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsnArg 361
      |||||:::
Db      1035 ---CTTGGTCCGGCTACTTCACTGCCATCCAGAACTTCAGGAAAAAGTGGAGGAA 1088
      |||||:::
QY      362 LeuPro-----ProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLys 378
      |||||:::
Db      1089 GTGGTCCCGCGGTGCGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1148
      |||||:::
QY      379 CysGlyAspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysVal 398
      |||||:::

```



[illegible]

RESULT 12

```

RES001 12
US-09-633-739--1
; Sequence 1, Application US/09633739
; Patent No. 6635447
; GENERAL INFORMATION:
; APPLICANT: Conneely, Orla M
; APPLICANT: Headon, Denis R.
; APPLICANT: O'Malley, Bert W.
; APPLICANT: May, Gregory S.
; TITLE OF INVENTION: PRODUCTION OF RECOMBINANT LACTOFERRIN
; TITLE OF INVENTION: AND LACTOFERRIN POLYPEPTIDES USING cDNA SEQUENCES IN VARIOUS
; TITLE OF INVENTION: ORGANISMS

```







Qy	188	LeuCysArgGlyAspSerSerGlyGluGlyValCysAspLysSerProLeuGluArgTyr	207
Db	591	CTGTGTGGGG-----ACAGGGGAAAAAATAATGTCTCTCTCCAGGAACCGTAC	644
Qy	208	TyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGlyAspValAlaPheVal	227
Db	645	TTTACGTACTCTGTGTGCTTCAAGTGTCTGAGAGACGGGCTGAGACGTGGCTTTATC	704
Qy	228	LysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuProSerTrpGlyGlnAla	247
Db	705	AGACAGAGCACAGTGTGGAGGCTCTCAGACGAGCT-----	743
Qy	248	LeuLeuSerGlnAspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGlu	267
Db	744	--GAAAGGACGAGTAGTAGTTACTCTGCCAGACAACACTCGGAAGCCAGTGGACAAG	800
Qy	268	TrpArgGlnCysHisLeuAlaArgValProAlaHisAlaValValValArgAla---	286
Db	801	TTCAAGACTGCCATCTGGCCGGGTCTCTCTCATCCGTTGTGGCACAAGCTGTGAAT	860
Qy	287	ThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGlnArgLeuPheSerHisGlu	306
Db	861	GGCAAGGAGATGCCATCTGAATCTTCTCGCCAGGCACAGGAAAGTTTGGAAAGGAC	920
Qy	307	GlySer---SerPheGlnMetPheSerSerGluAlaTyrGlyGlnLysAspLeuLeuPhe	325
Db	921	AAGTCACCGAAATTCAGACTCTTTGGCTCCCTAGT---GGGCAGAAAGATCTGCTGTC	977
Qy	326	LysAspSerThr-----SerGluLeuValPro---IleAlaThrGlnThrTyrGlu	341
Db	978	AAGACTCTGCATTGGGTTTTCAGGGTGCCCGCAGGATAGATTCTGGGCTGTAC--	1034
Qy	342	AlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsnArg	361
Db	1035	-----CTTGGCTCCGGCTACTTCACATGCCATCCAGAACTTGAGGAAAGTGAAGGAA	1088
Qy	362	LeuPro-----ProTyrLeuArgTyrCysValLeuSerThrProGluIleGlnLys	378
Db	1089	GTGGCTCCCGCGGTGCGGGTCTGTGTGTGGTGGCGGAGCAGAGCTGCGCAAG	1148
Qy	379	CysGlyAspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysVal	398
Db	1149	TGTAACCACTGGAGT-----GGCTTGAGCGAAGCAGCGCTGCTGCTCC	1193
Qy	399	SerAlaLysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaVal	418
Db	1194	TCGGCTCCACACAGAGGACTGCATCCCTCGTGTGAAAGGAGAAGCTGATGCCATG	1253
Qy	419	ThrLeuSerGlyGluAspIleTyrThAlaGlyLysLysTyrGlyLeuValProAlaAla	438
Db	1254	AGTTTGATGGAGGATATGTACATGCAGGC---AAATGTGGTTGGTGCTCTGCTG	1310
Qy	439	GlyGluHisTyrAlaProGluAspSerSerAsn-----	449
Db	1311	GCAGAGAATACAAATCCCAACAAAGCAGTGACCCCTGATCTTAACCTGTGTGATAGACCT	1370
Qy	450	-----SerTyrTyrValValAlaValAlaValArgAspSerSerHisAlaPheThrLeu	467
Db	1371	GTGGAAGGATATCTTGTGTGGCGGTGGTTAGGAGA---TCAGACACTAGCCTTACCTGG	1427
Qy	468	AspGluLeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAlaGlyTrpAsp	487
Db	1428	AACTCTGTGAAGGCAAGAGTCTCTGCCACACCGCGGTGACAGGACTGCAGGCTGGAAT	1487
Qy	488	ValProValGlyAlaLeuIleGlnArgGlyPheIleArgProLysAspCysAspValLeu	507
Db	1488	ATCCCCATGGGCTGCTC-----TTCAACCAGACGGGCTCTCTC-----	1526
Qy	508	ThrAlaValSerGluPhePheAsnAlaSerCysValProValAlaAsnProLysAsnTyr	527
Db	1527	---AAATTTGATGAATATTTCAAGTCAAGGCTGTGCCCTGGGCTGCAGCCGAGA----	1577

```

RESULT 14
US-08-655-640-1
; Sequence 1, Application US/08655640
; Patent No. 5948613
; GENERAL INFORMATION:
; APPLICANT: Teng, Christina
; APPLICANT: Panella, Timothy J.
; TITLE OF INVENTION: HUMAN LACTOFERRIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0. Version #1.25

```

; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/655,640  
 ; FILING DATE: 30-MAY-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/992,538  
 ; FILING DATE: December 17, 1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: SCOTT, WATSON T.  
 ; REGISTRATION NUMBER: 26,581  
 ; REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 861-3000  
 ; TELEFAX: (202) 822-0944  
 ; TELEX: 6714627 CUSH  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2117 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 1..2117  
 ; US-08-655-640-1

## Alignment Scores:

Pred. No.: 1,46e-124 Length: 2117  
 Score: 1261.00 Matches: 299  
 Percent Similarity: 55.16% Conservative: 118  
 Best Local Similarity: 39.55% Mismatches: 236  
 Query Match: 32.04% Indels: 104  
 DB: 2 Gaps: 28

US-10-049-957-4 (1-738) x US-08-655-640-1 (1-2117)

QY 3 GlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGlyMetGlu 22  
 DB 29 GGGCCCTCGGACTGTCTCTGGCTG-----GCCGTAGGAGA-AGGAGT 69  
 QY 23 ValArgTrpCysAlaThrSerArgProGluGlnHisCysGlyAsnMetSerGluAla 42  
 DB 70 GTTCAGTGTGGCCGTATATCCCAACCCGAGGCCACAAAATGCTTCCAAATGGCAAGGAAT 129  
 QY 43 PheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAlaAspHis 62  
 DB 130 ATGAGAAAGTGGT---GGCCCTCTGTCTCAGCTGTGATTAAGAGAGACTCCCCATCCAG 186  
 QY 63 CysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGlyAlaIle 82  
 DB 187 TGTATCCAGGCCATTGCGGAAACACAGGCCGATGCTGTGACCCCTTGATGGTGTTCATA 246  
 QY 83 TyrGluAlaGly---LysGluHisGlyLeuLysProValValGlyValGlyValTyr----- 99  
 DB 247 TAGGAGGAGGCGTGGCCCTTACAAACTGCGACCTGTAGCGCGGGAAGTCTACGGGACC 306  
 QY 100 AspGlnGluValGlyThrSerTyrTrpAlaValAlaValValArgSerSerHisVal 119  
 DB 307 GAAAGACACCCAGCACTACTATATATGCGTGGCTGTGTGTAAGAGGCGGCGACCTTT 366  
 QY 120 ThrIleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGly 139  
 DB 367 CAGCTGAACCAATGCAAGTCTGAAGTCTGCTGCCACACAGGCCCTTCGACGGACCGCTGGA 426  
 QY 140 TrpAsnValProValGlyTyrLeu----- 147  
 DB 427 TGAATGTCCCTATAGGACACTTCGTCATTCTTGAATTGACGGGTCCACTGAGCCC 486  
 QY 148 ValGluSerGlyArgLeuSerValMetGlyCysAspValLeuLysAlaValSerAspTyr 167  
 DB 487 ATTGAGCA----- 1503  
 QY 168 PheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSerGluSerLeuCysArg 187

DB 511 TTCTCAGCCAGCTGTGTCTCCGGTGCAGATAAAGGACAGTTC---CCCACTCTGTGTGCG 567  
 QY 188 LeuCysArgGlyAspSerSerGlyGluGlyValCysAspLysSerProLeuGluArgTyr 207  
 DB 568 CTGTGTGCGGG-----ACAGGGAAACAAMATGTCCCTCTCTCCCCAGGAACCGTAC 621  
 QY 208 TyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGlyAspValAlaPheVal 227  
 DB 622 TTCAGCTACTCTGGTGCCTTCAAGTGTCTGAGACGCGGGCTGAGACGCTGGCTTTTATC 681  
 QY 228 LysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuProSerTrpGlyGlnAla 247  
 DB 682 AGAGAGACACAGTGTGTTGAGGACCTGTGCAGCAGGCT----- 720  
 QY 248 LeuLeuSerGlnAspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGlu 267  
 DB 721 ---GAAAGGACAGTATGATTACTCTGCCAGACAACACTCGAAGCAGTGGACAAG 777  
 QY 268 TrpArgGlnCysHisLeuAlaArgValProAlaHisAlaValValValArgAla---Asp 286  
 DB 778 TTCAAAGACTGCCATCTGCCCGGGTCCCTTCTCATGCGCTTGTGACACAGTGTGAAT 837  
 QY 287 ThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGlnArgLeuPheSerHisGlu 306  
 DB 838 GGCAAGGAGGATGCCATCTTCTCCGCGCAGGCACAGGAAAAAGTTTGGAAAGGAC 897  
 QY 307 GlySer---SerPheGlnMetPheSerSerGluAlaTyrGlyGlnLysAspLeuPhe 325  
 DB 898 AAGTCACCGAAATTCAGCTCTTTGGCTCCCTAGT---GGGCAGAAAGATCTGCTGTTC 954  
 QY 326 LysAspSerThr-----SerGluLeuValPro---IleAlaThrGlnThrTyrGlu 341  
 DB 955 AAGGACTCTGCCATTGGGTTTTCGAGGGTGCCCGCAGGATAGATTCTGGCTGTAC--- 1011  
 QY 342 AlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsnArg 361  
 DB 1012 ---CTTGGCTCCGGCTACTTCACTGCATCCAGAACTTGAGGAAAAAGTGGAGGAA 1065  
 QY 362 LeuPro-----ProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLys 378  
 DB 1066 GTGGGTCCCGCGCGGT 1125  
 QY 379 CysGlyAspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysVal 398  
 DB 1126 TGTAAACAGTGGAGT-----GGCTTGAGCGAAGCGGACGCTGACCTGCTCC 1170  
 QY 399 SerAlaLysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaVal 418  
 DB 1171 TCGGCTCCACCAACAGAGACTGCATCGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1230  
 QY 419 ThrLeuSerGlyGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAla 438  
 DB 1231 AGTTTGGATGGAGGATATGTGTACACTGCAGGC---AAATGTGTTTGTGTGTGTGTGTGT 1287  
 QY 439 GlyGluHisTyrAlaProGluAspSerSerAsn----- 449  
 DB 1288 GCAGAGAACTACAAATCCCAACAGCAGTGCACCTGTATCTTAACCTGTGTGTGTGTGTGT 1347  
 QY 450 ---SerTyrTyrValValAlaValValArgArgAspSerSerHisAlaPheThrLeu 467  
 DB 1348 GTGGAAGGATATCTTGT 1404  
 QY 468 AspGluLeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAlaGlyTyrAsp 487  
 DB 1405 AACTCTGTGAAGGCAAGATCTTGCACACACCGCGGTGGACAGGACTCAGGCTCGAAT 1464  
 QY 488 ValProValGlyAlaLeuIleGlnArgGlyPheIleArgProLysAspCysAspValLeu 507  
 DB 1465 ATCCCCATGGGCTGCTC-----TTCAACACAGACGGGCTCCTGTC----- 1503  
 QY 508 ThrAlaValSerGluPhePheAsnAlaSerCysValProValProValProValProVal 527

Db 1504 ---AAATTTGATGAATAATTTCAAGTCAAAAGCTGTGCGCTGTGCTGACCCGAGCA----- 1554  
Qy 528 ProSerSerLeuGluValSerGlnPheAlaAlaCysAsnLeuGlnGlyArgAsnLysCysValGly 547  
Db 1555 ---TCTAATCTCTGTGCTGTGTATTTGGCGACGAGCGAGGGTGAAGATAAGTGGCTGCC 1611  
Qy 548 AenSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPheArgCysLeuValGluAsnAla 567  
Db 1612 AACACCAACGAGAGATACAGCTACAGCTGGGCTTTCCGGTGTGCTGGTGAAGATGCT 1671  
Qy 568 GlyAspValAlaPheValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSer 587  
Db 1672 CGAGACGTTGCAATTTGTGAAGATGCTCACTGTCTTGAGAACACACTGATGTAATTAACAAT 1731  
Qy 588 GluProTyrAlaAlaGluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAla 607  
Db 1732 GAGGCGATGGCTTAAGATTTGAAGTGGCAGACTTTGGCGCTGTGCTGTGCTGATGGCAAA 1791  
Qy 608 ArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAlaGlnIleProProHisAlaVal 627  
Db 1792 CGGAGCTGTGACTGAGCTAGAGCTGCCATCTTGGCATGGCCCGCCGAATCATGCCGTG 1851  
Qy 628 MetValArgProAspThrAsnIlePheThrValTyrGlyLeuLeuAspLysAlaGlnAsp 647  
Db 1852 GTGTCTCGATGGAT---AAGGTGAACGCTGAAACAGGTGTGTCTCCACCAACAGGCT 1908  
Qy 648 LeuPheGlyAspAspHisAsnLysAsnGly-----PhelysMetPhe 661  
Db 1909 AAATTTGGG-----AGAAATGGATCTGACTGCCCGGACAAAGTTTGTCTTATTC 1956  
Qy 662 AspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrValArgAlaVal 681  
Db 1957 CAGTCTGAAC-----AAAACTTCTGTTCAATGACCAACACTGAGTGTCTGCC 2007  
Qy 682 ProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAlaLeu 701  
Db 2008 AGACTCCATGGCAAAACAAATATGAAATAATTTGGGACCAACAGATATGTGCGAGGCAAT 2067  
Qy 702 GluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaPro 717  
Db 2068 ACTAATCTG-----AAAAAGTGTCAACCTCCCTCTCTGGAAGCCT 2109

## RESULT 15

US-08-655-640-3  
; Sequence 3, Application US/08655640  
; Patent No. 5948613  
; GENERAL INFORMATION:  
; APPLICANT: Teng, Christina  
; APPLICANT: Panella, Timothy J.  
; TITLE OF INVENTION: HUMAN LACTOFERRIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESSES:  
; ADDRESS: CUSHMAN, DARBY & CUSHMAN  
; STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR  
; CITY: WASHINGTON  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/655,640  
; FILING DATE: 30-MAY-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/992,538  
; FILING DATE: December 17, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SCOTT, WATSON T.  
; REGISTRATION NUMBER: 26,581

; REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2124 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..2124  
; US-08-655-640-3  
Alignment Scores:  
Pred. No.: 1.66e-124 Length: 2124  
Score: 1260.50 Matches: 301  
Percent Similarity: 54.93% Conservative: 117  
Best Local Similarity: 39.35% Mismatches: 233  
Query Match: 32.02% Indels: 111  
DB: 2 Gaps: 29  
US-10-049-957-4 (1-738) x US-08-655-640-3 (1-2124)  
Qy 3 GlyProSerGlyAlaLeuTyrPheLeuLeuAlaLeuArgThrValLeuGlyGlyMetGlu 22  
Db 35 GGGCCCTCGGACTGTGTCTGGCTG-----GCCGTAGGAGA-AGGAGT 75  
Qy 23 ValArgTyrCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSerGluAla 42  
Db 76 GTTCAGTGTGGCGGTATCCCAACCCGAGGCCACAAATGCTTCCAATGCCAAAGGAAT 135  
Qy 43 PheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAlaAspHis 62  
Db 136 ATGAGAAAGTGGCT---GGCCCTCTCTGACGTGCTATAAAGAGAGAGACTCCCCCATCCAG 192  
Qy 63 CysValGlnLeuIleAlaAlaGlnGluAlaAspAlaIleThrLeuAspGlyGlyAlaIle 82  
Db 193 TGTATCCAGGCCATTGGCGAAACACAGGCCCATGCTGTGACCCCTTGTATGGTGTTCATA 252  
Qy 83 TyrGluAlaGly---LysGluHisGlyLeuLysProValValGlyGlyValTyr----- 99  
Db 253 TACGAGGACGCTGGGCCCCCTACAACTGGAGCTGTAGCGGCGGAAGTCTACGGGACC 312  
Qy 100 AspGlnGluValGlyThrSerTyrTyrAlaValAlaValArgArgSerSerHisVal 119  
Db 313 GAAAGACAGCCACGAACTCACTATTATGCCGTGGCTGTGGTGAAGAGGGCGGAGCTTT 372  
Qy 120 ThrIleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGly 139  
Db 373 CAGCTGAACGAACTGCAAGGTCTGAAGTCTCTGCCACACAGGCCCTTCGACGAGGACCGTGA 432  
Qy 140 TrpAsnValProValGlyTyrLeu----- 147  
Db 433 TGGAAATGTCCCTATAGGACACTTCGTCCATTCTTGAATTGACGGGTCCACCTGAGCCC 492  
Qy 148 ValGluSerGlyArgLeuSerValMetGlyCysAspValLeuLysAlaValSerAspTyr 167  
Db 493 ATTGAGGCA-----GCTGTGGCCAGGTTTC 516  
Qy 168 PheGlyGlySerCysValProGlyAlaGlyGluThrSerTyrSerGluSerLeuCysArg 187  
Db 517 TTCTCAGCAGCTGTGTTCGCGGTGCAGATAAAGGACAGTTTC---CCAACTGTGTGTCGC 573  
Qy 188 LeuCysArgGlyAspSerSerGlyGlyValCysAspLysSerProLeuGluArgTyr 207  
Db 574 CTGTGTGGGGG-----ACAGGGGAAACAAATGTGCTTCTCTCCCGAGAACCGTAC 627  
Qy 208 TyrAspTyrSerGlyAlaPheArgCysLeuAlaGlyAlaGlyAspValAlaPheVal 227  
Db 628 TTCAGCTACTCTGTGTGCTTCAAGTGTCTGAGAGACGGGGCTGGAGAGCTGTGCTTTTATC 687

```
QY 228 LysHisSerThrValLeuGluAenThrAspGlyLysThrLeuProSerTrpGlyGlnAla 247
   ::::::::::::::::::::
Db 688 AGAGAGACACAGCTGTTGAGGACCTGTGACAGCAGGCT----- 726

QY 248 LeuLeuSerClnAspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGlu 267
   ::::::::::::::::::::
Db 727 ---GAAAGGACAGTGTGAGTTACTCTGCCACAGACACACTCGGAAGCCAGTGGACAAG 783

QY 268 TrpArgClnCysHisLeuAlaArgValProAlaHisAlaValValValAlaAla---Asp 286
   ::::::::::::::::::::
Db 784 TTCAAAGACTGCCATCTGCCCGGCTCCCTTCTCATGCCGTTGTGGCAGGAAGTGTGAAT 843

QY 287 ThrAspGlyGlyLeuLeuPheArgLeuLeuAenGluGlyClnArgLeuPheSerHisGlu 306
   ::::::::::::::::::::
Db 844 GGCAGGAGGATGCCATCTGGAATCTCTCCGACAGGCACAGGAAAAGTTTGGAAAGGAC 903

QY 307 GlySer---SerPheGlnMetPheSerSerGluAlaTyrglyClnLysAspLeuLeuPhe 325
   ::::::::::::::::::::
Db 904 AGTCAACCGAAATTCAGCTCTTTGGCTCCCTAGT---GGGCAGAAAGATCTGCTGTTTC 960

QY 326 LysAspSerThr-----SerGluLeuValPro---IleAlaThrGlnThrThrGlu 341
   ::::::::::::::::::::
Db 961 AGGACTCTGCCATTGGGTTTTCGAGGCTGCCCGCAGGATAGATTCTGGGCTGTAC--- 1017

QY 342 AlaTrpLeuGlyHisGluTyrluHisAlaMetLysGlyLeuLeuCysAspProAsnArg 361
   ::::::::::::::::::::
Db 1018 ---CTTGGCTCCGGCTACTTCACTGGCCATCCAGAACTTGAGGAAAAGTGAGGAGAA 1071

QY 362 LeuPro-----ProTyrluArgTrpCysValLeuSerThrProGluLeuGlnLys 378
   ::::::::::::::::::::
Db 1072 GTGGCTGCCCGGCGTGGGCTGGTGTGGTGGGCGGAGCAGGAGCTGGCCAAAG 1131

QY 379 CysGlyAspMetAlaValAlaPheArgArgGlnArgLeuLysProGluLeuGlnCysVal 398
   ::::::::::::::::::::
Db 1132 TGTAACCACTGGAGT-----GGCTTCAGCGAAGCGCGTGCCTCTCC 1176

QY 399 SerAlaLysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAlaVal 418
   ::::::::::::::::::::
Db 1177 TCGGCTCCACCAACAGAGACTGCATCGCCCTGGTGTGAAAGGAGAAGCTGATGCCATG 1236

QY 419 ThrLeuSerGlyClnAspIleTyrluAlaGlyLysTyrglyLeuValProAlaAla 438
   ::::::::::::::::::::
Db 1237 AGTTTGGATGGAGGATATGTGTACACTGCAGGC---AAATGTGTTTGGTGCCTGTCTG 1293

QY 439 GlyGluHisTyrluAlaProGluAspSerSerAsn----- 449
   ::::::::::::::::::::
Db 1294 GCAGAGAACTACAAATCCCAACAAAGCAGTGACCTGATCCTTAACGTGTGGATAGACCT 1353

QY 450 -----SerTyrluValValAlaValAlaValArgAspSerSerHisAlaPheThrLeu 467
   ::::::::::::::::::::
Db 1354 GTGGAAGGATATCTGTGTGGGGTGTGTAGGAGA---TCAGACACTAGCCTTACCTGG 1410

QY 468 AspGluLeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAlaGlyTrpAsp 487
   ::::::::::::::::::::
Db 1411 AACTCTGTGAAGGGAAGAGTCCCTCCACACCGCGCTGGACAGGACTCGAGCTGGAAT 1470

QY 488 ValProValGlyAlaLeuLeuGlnArgGlyPheIleArgProLysAspCysAspValLeu 507
   ::::::::::::::::::::
Db 1471 ATCCCCATGGGCTGTCTC-----TTCAACACAGCGGCTCCTGC----- 1509

QY 508 ThrAlaValSerGluPhePheAsnAlaSerCysValProValAsnAsnProLysAsnTyrlu 527
   ::::::::::::::::::::
Db 1510 ---AAATTTGATGAATATTTCAAGCAAGCTGTGCCCTGGGTCTGACCCGAGA----- 1560

QY 528 ProSerSerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGly 547
   ::::::::::::::::::::
Db 1561 ---TCTAATCTCTGTCTGTGTATTTGGCGACGAGCGGTGAGAAATAGTGCCTGCC 1617

QY 548 AsnSerGlnGluArgTyrluTyrluArgGlyValAlaPheArgCysLeuValGluAsnAla 567
   ::::::::::::::::::::
Db 1618 AACAGCAACAGAGAGATACAGGCTACACTGGGGCTTTCCGGTGCCTGGCTGAGAATGCT 1677
```

```
QY 568 GlyAspValAlaPheValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSer 587
   ::::::::::::::::::::
Db 1678 GGAGACGTTGCTATTTGTGAAGATGTCTACTGTCTTCGACAGACACTGATGGAATAACAAT 1737

QY 588 GluProTrpAlaAlaGluLeuArgSerGluAspTyrluLeuLeuCysProAsnGlyAla 607
   ::::::::::::::::::::
Db 1738 GAGGCATGGGCTAAGATTGAGCTGGCAGACTTTCCTGCTGTCTGCTCGATGCGAAA 1797

QY 608 ArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAlaGlnIleProProHisAlaVal 627
   ::::::::::::::::::::
Db 1798 CGGAAGCCTGTGCTAGCTAGAGCTGCCATCTTGCCTAGCCCGAATCATGCGGTG 1857

QY 628 MetValArgProAspThrAsnIlePheThrValTyrluLeuAspLysAlaGlnAsp 647
   ::::::::::::::::::::
Db 1858 GTGTCTCGGATGGAT---AAGGTGGAACGCTGAAACACAGGTGTGTCTCCACCAACAGGCT 1914

QY 648 LeuPheGlyAspAspHisAsnLysAsnGly-----PheLysMetPhe 661
   ::::::::::::::::::::
Db 1915 AAATTTGGG-----AGAAATGGATCTGACTGCCCGGACAAAGTTTTCCTATTTC 1962

QY 662 AspSerSerAsnTyrluHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAlaVal 681
   ::::::::::::::::::::
Db 1963 CAGTCTGAAACC-----AAAAACCTTCTGTTCATGACAACTGAGTGTCTGGCC 2013

QY 682 ProValGlyGluLysThrThrTyrluArgGlyTrpLeuGlyLeuAspTyrluAlaAlaLeu 701
   ::::::::::::::::::::
Db 2014 AGACTCCATGGCAAAACACATATGAAAAATATTTGGGACCACTGATGTGCGAGCAAT 2073

QY 702 GluGlyMetSerSerGlnGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaProLeu 721
   ::::::::::::::::::::
Db 2074 ACTAATCTG-----AAAAAGTGTCTCA-----ACCTCCCCCCTC 2106

QY 722 Leu 722
   |||
Db 2107 CTG 2109
```

Search completed: May 16, 2004, 23:37:13  
Job time : 231 secs

**This Page Blank (usp10)**



Y

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 16, 2004, 19:30:57 ; Search time 4867 Seconds  
(without alignments)  
4528.105 Million cell updates/sec

Title: US-10-049-957-4  
Perfect score: 3936  
Sequence: 1 MRGPGALWLLALRTVLGG.....APLLPLLPAALRLLPPAL 738

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPIO.spool/US1004957/runat\_14052004\_100250\_1269/app\_query.fasta\_1.903  
-DB=EST -Qfmt=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US1004957 @CGN 1.1 1906 @runat\_14052004\_100250\_1269 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pin.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	DB	ID	Description
1	3300	83.8	4055	11	AK036444	Mus muscu AK036444 Mus muscu
2	2375	60.3	1734	29	AY420522	Homo sapi AY420522 Homo sapi
3	2060	52.3	1734	29	AY420524	Mus muscu AY420524 Mus muscu
4	1435	36.5	1734	29	AY420523	Pan trogl AY420523 Pan trogl
5	1245	31.6	1023	13	BUI185213	AGENCOURT BUI185213 AGENCOURT
6	1244	31.6	937	13	BQ29045	AGENCOURT BQ29045 AGENCOURT
7	1237	31.4	998	13	BUI168950	AGENCOURT BUI168950 AGENCOURT
8	1224	31.1	819	10	BE274850	Mus muscu BE274850 Mus muscu
9	1210	30.7	2313	11	BC058216	Mus muscu BC058216 Mus muscu
10	1202	30.5	2314	11	BC020295	Mus muscu BC020295 Mus muscu
11	1202	30.5	2316	11	AK085754	Mus muscu AK085754 Mus muscu
12	1202	30.5	2332	11	BC058218	Mus muscu BC058218 Mus muscu
13	1201.5	30.5	2249	11	AY327504	Rattus no AY327504 Rattus no
14	1199.5	30.5	2296	11	BC008530	Mus muscu BC008530 Mus muscu
15	1195.5	30.4	2744	11	AK036491	Mus muscu AK036491 Mus muscu
16	1152	29.3	918	10	BE727854	Mus muscu BE727854 Mus muscu
17	1144.5	29.1	2886	11	AY325214	Rattus no AY325214 Rattus no
18	1144.5	29.1	2986	11	AY325241	Rattus no AY325241 Rattus no
19	1136.5	28.9	1054	10	BE728795	Mus muscu BE728795 Mus muscu
20	1131	28.7	808	10	BE728263	Homo sapi BE728263 Homo sapi
21	1122	28.5	2153	11	AF118063	Mus muscu AF118063 Mus muscu
22	1121	28.5	2989	11	AY325230	Rattus no AY325230 Rattus no
23	1089.5	27.7	2277	11	AK005035	Mus muscu AK005035 Mus muscu
24	1087	27.6	1099	13	BUS42074	AGENCOURT BUS42074 AGENCOURT
25	1081	27.5	666	10	BE275242	Mus muscu BE275242 Mus muscu
26	1057	26.9	748	10	BE408902	AGENCOURT BE408902 AGENCOURT
27	1046	26.6	707	9	AI121147	us89H07.Y AI121147 us89H07.Y
28	1045	26.5	798	12	BG763242	Mus muscu BG763242 Mus muscu
29	1027.5	26.1	1043	12	BG333821	Mus muscu BG333821 Mus muscu
30	1016	25.8	852	12	B1696630	Mus muscu B1696630 Mus muscu
31	1008	25.6	712	10	BE385618	Mus muscu BE385618 Mus muscu
32	1002.5	25.5	659	10	BE390516	Mus muscu BE390516 Mus muscu
33	1001	25.4	745	10	BE408846	Mus muscu BE408846 Mus muscu
34	999	25.4	660	10	BE408717	Mus muscu BE408717 Mus muscu
35	997.5	25.3	692	9	AI479814	tm69f12.x AI479814 tm69f12.x
36	989.5	25.1	755	10	BE410394	Mus muscu BE410394 Mus muscu
37	978.5	24.9	871	13	BQ686877	AGENCOURT BQ686877 AGENCOURT
38	969.5	24.6	729	10	BE277294	Mus muscu BE277294 Mus muscu
39	969	24.6	611	10	BE386317	Mus muscu BE386317 Mus muscu
40	931	23.7	823	13	BUI37676	Mus muscu BUI37676 Mus muscu
41	899.5	22.9	860	14	CF250645	Mus muscu CF250645 Mus muscu
42	881.5	22.4	848	12	BG478400	Mus muscu BG478400 Mus muscu
43	881	22.4	663	10	BB629343	Mus muscu BB629343 Mus muscu
44	879.5	22.3	674	10	BE729755	Mus muscu BE729755 Mus muscu
45	868	22.1	701	10	BE387819	Mus muscu BE387819 Mus muscu

ALIGNMENTS

RESULT 1  
AK036444  
LOCUS  
DEFINITION  
AK036444  
Mus musculus adult male bone cDNA, RIKEN full-length enriched library, clone:9830108P05 product:antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5, full insert sequence.  
ACCESSION  
AK036444  
VERSION  
AK036444.1  
GI:26331403  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
4055 bp  
mRNA  
linear  
HTC 19-SEP-2003

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 PUBMED 12040555

6 (bases 1 to 4055)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/  
 URL: http://location.qualifiers

FEATURES  
 source 1..4055  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"

/db\_xref="PANTOM,DB:9830108P05"  
 /db\_xref="MGI:240871"  
 /db\_xref="taxon:10090"  
 /clone="9830108P05"  
 /sex="male"  
 /tissue type="bone"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="adult"

53..2289  
 /note="unnamed protein product; antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5 (MGD|MGI:1353421, GB|NM\_013900, evidence: BLASTN, 99% match=4055)  
 putative"

/codon\_start=1  
 /protein\_id="BAC29432.1"  
 /db\_xref="GI:26331404"  
 /translation="MRLSVTVFLLLSLRTVVCVMEVQWCTISDAEQKQKDMSEAFQ GAGIRPSLLCVQNSADHCVOLIKEQKADAITLDGAIYEAGKEHGLKPVVGVVVDQ IGTSYAVAVVRNSNTINTLKGVSCHTGINRTVGNVPVGVLSHLSVMGCDV LKAVDVGSGCVPTGETSHSESLCLCRGSSGHNVCDKSPLERYDYSGARCLA EGADGAVVRHSTVLENTDNTLPSWKSLSMSEDFQLLCRGSRADITTEWRCHLAKV PAHAAVVRGMDGLIFQLNEGQLLFSHEDSFPFSSKAYSQKLLFKDSTLELVP IATQCVSAESPEHCMEQIQAGTDAVTLRGEDIVRAGKVYIPAAAGLYAEEDRSNS YFVAVARROSSYSFTLDELGRKBSCHPYLGSFAGMEVPIGSLIQRGFIRPKDCDLVT AVSQFFNASCVPVNNPNKYPFALCALCVGDEGRKNCVGSQERYYSFGAPRLVEH AGDVAFVHTTTFENTGNHPNPNPASHLRWDYELLCPNGARAEVYDQFQANLQMP5 HAVMVRPTNIFTYVGLLDKAQDLFGDDHNNKQFMDFSKYSQDLDFKDATVRAVP VREKTTVLDMLGPDYVVALEGLMQSCSGAGAAVQRPVPLALLLLTAAAGLLPVL"

# ORIGIN

Alignment Scores:  
 Pred. No.: 0 Length: 4055  
 Score: 3300.00 Matches: 618  
 Percent Similarity: 89.84% Conservative: 45  
 Best Local Similarity: 83.74% Mismatches: 75  
 Query Match: 83.84% Indels: 0  
 Ds: 11 Gaps: 0

US-10-049-957-4 (1-738) x AK036444 (1-4055)  
 QY 1 MetArgGlyPProSerGlyAlaLeuTriLeuLeuAlaLeuArgThrValLeuGlyGly 20  
 DB 53 ATGAGGCTCTCGAGCGTGACTTTTGGCTATCTCTGCTCCCTCGGCACATGTGCTGTGTG 112  
 QY 21 MetGluValArgTTPCyseAlaThrSerAspProGluGlnHisLysCyseGlyAsnMetSer 40  
 DB 113 ATGGAGGTGCAGTGTGTATCCATCTCAGACGAGCAGCAGAGCAAGTGCAGGACATGAGC 172  
 QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
 DB 173 GAGGCTTCCAGGAGGAGTGGCATTCGCTCTTCTCTGCTCCGTCAGGGCAACTCCGCT 232  
 QY 61 AsphHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80  
 DB 233 GACCACTGTGTCCAGCTCATCAAGAACAAAGACAGATGCATCACCTGGATGAGGG 292  
 QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyValTyrAsp 100  
 DB 293 GCCATCTATGAGGAGGAGGAGGAGCAGCGGCTGAAGCCAGTGTGGGGGAAAGTCTATGAC 352  
 QY 101 GlnGluValGlyThrSerTyrAlaValAlaValAlaValArgSerSerHisValThr 120  
 DB 353 CAAGACATTTGGGACTTCTATTATGCGGTGCTGTGGTTCAGGAGAAATTCATGTTACC 412  
 QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140  
 DB 413 ATCAACACCTGAAGGGGTCAAGTCTCCACACAGAGCATTAACCGACATGTGGGCTGG 472  
 QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
 DB 473 AACGTGCTGTCGGTTACCTCGTAGAGAGCGGCCCATCTGTGTCAGTGTGGGCTGTGTGTG 532



```
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180
Db 533 CTCAAGCCGTTGGTGATTTATTTGGAGGAGCTGTGCTCCCTGGAACAGAGAAACGAGC 592
QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200
Db 593 CATTCGAGTCCCTCTGTCGCTCTGCGGTGGGAGCTCTTCTGGGCAACAATGTGTGTGAC 652
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly 220
Db 653 AAGAGTCCCTAGAGAGATCTACGACTACAGTGGAGCCTTCCGGTGCCTGGCGGAAGGA 712
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240
Db 713 GCCGGTGACGTGGCTTCTGGTGAAGCAGCAGCACAGTGTGGAATAATCTGATGAAACACC 772
QY 241 LeuProSerTrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGly 260
Db 773 CTGCTCTTCTGGGGCAAGTCCCTGATGTACAGAGACTTCCAGCTACTATGACGGGATGGC 832
QY 261 SerArgAlaAspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAla 280
Db 833 AGCCGAGCCACATCACTAGTGGAGACGTGGCCACTGGCCAAAGGTGCTGCTCATGCT 892
QY 281 ValValAlaArgAlaAspThrAspGlyGlyLeuIlePheArgLeuLeuAsnGluGlyGln 300
Db 893 GTGGTGGTCAAGGGGTGACATGATGGCGGTCTCATATTCCAATGCTCAACGAGGCCAG 952
QY 301 ArgLeuPheSerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGln 320
Db 953 CTTCTGTTCAGCCATGAAGCAGCAGCTTCCAGATGTTCCAGTCCCAAGCCTACAGCCAG 1012
QY 321 LysAspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThrTyr 340
Db 1013 AAGAATCTGTGTTCAAAAGACTCCACTTGGAGCTTGTGCCCATTCACACAGAACTAT 1072
QY 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeuLeuCysAspProAsn 360
Db 1073 GAGGCTGTGGTGGCCAGGATACCTGCGGCCATGAGGGGCTCTCTGTGATCCCAAC 1132
QY 361 ArgLeuProProTyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGly 380
Db 1133 CGGCTGCCCCACTACCTGCGCTGGTGTGCTGTCCAGCGCCGAGATCCAGAAGTGTGA 1192
QY 381 AspMetAlaValAlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAla 400
Db 1193 GATATGGCTGTGGCTTTCAGCCGCCAGATCTCAAGCCGGAATTCAGTGTGTGTCGGCC 1252
QY 401 LysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAlaValThrLeu 420
Db 1253 GAGTCCCTCGAGCACTGCATGGAGCAGATCCAGGCTGGGCACACTGACGCTGTGATCTTG 1312
QY 421 SerGlyGluAspTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGlu 440
Db 1313 AGGGGGCAGACATTTACAGGGCAGGAAGGTGTACGGCTGTGTTCCGGCGGCGGGGAG 1372
QY 441 HisTyrAlaProGluAspSerSerAsnSerTyrTyrValValAlaValValArgArgAsp 460
Db 1373 CTGTATGCTGTAGAGACAGGACCAATTCCTACTTCTGTGTGTGCTGTGCGACAGAGGAC 1432
QY 461 SerSerHisAlaPheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPhe 480
Db 1433 AGTCTCTCTCTTCCCTGAGCAGAGCTTCCGGCAGAGGCTTCTGCCACCCCTACTTG 1492
QY 481 GlySerProAlaGlyTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArg 500
Db 1493 GGCAGCCCGAGGGCTGGAGGTCGCCATCGGCTCCCTCATCCAGCGGGCTTTCATCCGG 1552
QY 501 ProLysAspCysAspValLeuThrAlaValSerGluPheAsnAlaSerCysValPro 520
Db 1553 CCCAAGGACTGTGATGTCTTCACAGCGGTGAGCCAGTCTTCAATGCCAGCTCGTGCCT 1612
```

---

```
QY 521 ValAsnAsnProLysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGln 540
Db 1613 GTCAACAACCCCTAAGAACTACCTTCCGCACTATGTGCGCTCTGCGTGGGAGACGAGAAG 1672
QY 541 GlyArgAsnLysCysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyAlaPhe 560
Db 1673 GGCCCGCAACAATGTGTGGGAGCAGCCAGGAGATACTACGGCTACAGCGGGGCTTC 1732
QY 561 ArgCysLeuValGluAsnAlaGlyAspValAlaPheValArgHisThrThrValPheAsp 580
Db 1733 AGTGTGCTTGTGGAGCATGACGGGACGTGGCTTCTCAAGCACACGACTGTCTTTGAG 1792
QY 581 AsnThrAsnGlyHisAsnSerGluProTrpAlaAlaGluLeuArgSerGluAspTyrGlu 600
Db 1793 AACACAATGGTCACAATCTGAGCCTTGGGCTTCTCACCTCAGGTGGCAAGACTATGAA 1852
QY 601 LeuLeuCysProAsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAla 620
Db 1853 CTACTGTGCCCCAATGGGGCAGCGGCTGAGGTAGACCACTTCCAAGCTTGGCAACCTGGCA 1912
QY 621 GlnIleProProHisAlaValMetValArgProAspThrAsnIlePheThrValTyrGly 640
Db 1913 CAATGCGCATCCACGCTGTATGTCCGTCCAGACACCAATCTTCACTGTGTATGGA 1972
QY 641 LeuLeuAspLysAlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGlyPheLysMet 660
Db 1973 CTTCTGGACAAGGCCCAGGACCTGTTTGGAGAGCACCATAACAAGAACGTTTCCAAATG 2032
QY 661 PheAspSerSerAsnTyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValArgAla 680
Db 2033 TTTGACTCTCTCCAAATATCACAGCCAAAGACCTGCTTTTCAAGATGCTCAGTCCGAGCG 2092
QY 681 ValProValGlyGluLysThrThrTyrArgGlyTyrLeuGlyLeuAspTyrValAlaAla 700
Db 2093 GTGCAGTCCGGAGAAAACCACTACCTGGACTGGCTGGGCTCCTGACTATGTGTGCG 2152
QY 701 LeuGluGlyMetSerSerGlnCysSerGlyAlaAlaAlaProAlaProGlyAlaPro 720
Db 2153 CTGGAGGGAGTGTGTCTCAGCAGTGTCTCCGGTGCAGGGCGCGGTGCAGCGAGTCCCC 2212
QY 721 LeuLeuProLeuLeuLeuProAlaLeuAlaAlaArgLeuLeuProProAlaLeu 738
Db 2213 CTGCTGGCCCTGCTCCTGTGCTGACCCCTGGCTGCAGGCTCCTCTCCCTCGCTTCTC 2266

RESULT 2
AY420522 1734 bp DNA linear GSS 17-DEC-2003
LOCUS Homo sapiens MF12 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY420522
VERSION AY420522.1 GI:39776479
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1734)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1734)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
```

COMMENT This sequence as made by sequencing genomic exons and ordering them based on alignment.

## FEATURES

Location/Qualifiers

source

1..1734  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>1734  
/gene="MF12"  
/locus\_tag="HCM7252"

## gene

## ORIGIN

## Alignment Scores:

Pred. No.: 2,63e-238 Length: 1734  
Score: 2375.00 Matches: 463  
Percent Similarity: 80.10% Conservative: 0  
Best Local Similarity: 80.10% Mismatches: 115  
Query Match: 60.34% Indels: 0  
DB: 29 Gaps: 0

US-10-049-957-4 (1-738) x AY420522 (1-1734)

QY 69 AlaGlnGluAlaAspAlaIleThrLeuAspGlyGlyAlaIleTyrGluAlaGlyLysGlu 88  
DB 1 GCCCAGGAGGCTCACGCCATCACTCTGGATGGAGGAGCCATCTATGAGCGGGAAAGGAG 60  
QY 89 HisGlyLeuLysProValValGlyGluValTyrAspGlnGluValGlyThrSerTyrTyr 108  
DB 61 CACGGCCCTGAAGCCGGTGGTGGCAAGTGTACGATCAAGNNNNNNNNNNNNNNNNNN 120  
QY 109 AlaValAlaValAlaArgSerHisValThrIleAspThrLeuLysGlyValLys 128  
DB 121 NNN 180  
QY 129 SerCysHisThrGlyIleAsnArgThrValGlyTyrAsnValProValGlyTyrLeuVal 148  
DB 181 NNN 240  
QY 149 GluSerGlyArgLeuSerValMetGlyCysAspValLeuLysAlaValSerAspTyrPhe 168  
DB 241 NNN 300  
QY 169 GlyGlySerCysValProGlyAlaGlyGluThrSerTyrSerGluSerLeuCysArgLeu 188  
DB 301 NNN 360  
QY 189 CysA-gGlyAspSerSerGlyGluGlyValCysAspLysSerProLeuGluArgTyrTyr 208  
DB 361 NNN 420  
QY 209 AspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGlyAspValAlaPheValLys 228  
DB 421 NNN 480  
QY 229 HisSerThrValLeuGluAsnThrAspGlyLysThrLeuProSerTyrGlyGlnAlaLeu 248  
DB 481 CACAGCAGCGTACTGGAGAACACGATGGGAAGACGCTTCCCTCTGGGGCCAGCCCTG 540  
QY 249 LeuSerGlnAspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGluTrp 268  
DB 541 CTGTACAGGACTTCAGCTGCTGTCGGGATGTGTAGCCGGCCGATGTCTACCCGAGTGG 600  
QY 269 ArgGlnCysHisLeuAlaArgValProAlaHisAlaValValAlaArgAlaAspThrAsp 288  
DB 601 AGGCAGTGCATCTGGCCGGGTGCTGCTACGCGCGTGGTGGTCCGGCCGACACAGAT 660  
QY 289 GlyGlyLeuIlePheArgLeuLeuAsnGluGlyGlnArgLeuPheSerHisGlyGlySer 308  
DB 661 GGGGGCCCTCATCTTCGGGCTGTCTACAGNAGCCAGCGTCTGTTTACGCCACGAGGGCAGC 720  
QY 309 SerPheGlnMetPheSerSerGluAlaTyrGlyGlnLysAspLeuPheLysAspSer 328  
DB 721 AGCTTCCAGATGTTCAAGTCTCTGAGGCTTATGGCCAGAGGATCTACTCTTCAAGACTCT 780

QY 329 ThrSerGluLeuValProIleAlaThrThrTyrGlnAlaIleThrLeuGlyHisGluTyr 348  
DB 781 ACCTCGGAGCTTGTGCCCATGCCACACAGACCTATGAGGGGTGGCTGGGCATGAGTAC 840  
QY 349 LeuHisAlaMetLysGlyLeuLeuCysAspProAsnArgLeuProProTyrLeuArgTrp 368  
DB 841 CTGCAGCCCATGAAGGGTCTGCTCTGTGACCCCAACCGGCTGCCCCCTACTCTGGCTGG 900  
QY 369 CysValLeuSerThrProGluIleGlnLysCysGlyAspMetAlaValAlaPheArgArg 388  
DB 901 TGTGTGCTCTCCACTCCCGAGATCCAGAAGTGTGGAGACATGGCGGTGGCTTCCGCCGG 960  
QY 389 GlnArgLeuLysProGluIleGlnCysValSerAlaLysSerProGlnHisCysMetGlu 408  
DB 961 CAGCGCTCAAGCCAGAGATCCAGTGGTGTGCGCAAGTCCCCCAACACTGCATGTGAG 1020  
QY 409 ArgIleGlnAlaGluGlnValAspAlaValThrLeuSerGlyGluAspIleTyrThrAla 428  
DB 1021 CGGATCCAGGCTGAGCAGGTCGACGCTGTGACCTGAGTGGCGAGGACATTTACACGGCG 1080  
QY 429 GlyLysLysTyrGlyLeuValProAlaAlaGlyGluHisTyrAlaProGluAspSerSer 448  
DB 1081 GGGAAAGACGTACGGCTGTTCCGCGACGCCGGGAGCACTATGCCCGCGAAGACAGCAGC 1140  
QY 449 AsnSerTyrTyrValValAlaValAlaValArgAspSerSerHisAlaPheThrLeuAsp 468  
DB 1141 AACTCGTACTACGTGGTGGCGCTGGTGGAGCGGAGCAGCTCCACGCTTCCCTTGGAT 1200  
QY 469 GluLeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAlaGlyTyrAspVal 488  
DB 1201 GAGCTTCGGGCAAGCGCTCTGCGCACGCCGGTTTCGCGACGCCCTGCAGGCTGGATGTC 1260  
QY 489 ProValGlyAlaLeuIleGlnArgGlyPheIleArgProLysAspCysAspValLeuThr 508  
DB 1261 CCCGTGGGTGGCTTATTACAGAGAGGCTTCAATCGCGCCCAAGAGAGTGTGAGCTCTCACA 1320  
QY 509 AlaValSerGluPhePheAsnAlaSerCysValProValAlaAsnProLysAsnTyrPro 528  
DB 1321 CGAGTGAGCGAGTGTCTTCAATGCCAGCTGCGTGGCGGGAACAACCCCAAGACTACCCC 1380  
QY 529 SerSerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGlyAsn 548  
DB 1381 TCCTCGCTGTGTGCACGTGTGCGGGGACGAGCAGCGCGCAACAAGTGTGGGCAAC 1440  
QY 549 SerGlnGluArgTyrTyrGlyArgGlyAlaPheArgCysLeuValGluAsnAlaGly 568  
DB 1441 AGCCAGGAGCGGTATTACGGCTACCGCGCGCTTCAAGTGTGGTGGAGATGCGGGT 1500  
QY 569 AspValAlaPheValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSerGlu 588  
DB 1501 GACGTGTGCTTCCTGTCAGGCACACACCGCTTTTGACACACACACCGGCCCAATTCGAG 1560  
QY 589 ProTrpAlaAlaGluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAlaArg 608  
DB 1561 CCCTGGGCTGTGAGCTCAGGTACAGGACTATGAAGTGTGCTGCGCCCAACCGGGCCCGA 1620  
QY 609 AlaGluValSerGlnPheAlaAlaCysAsnLeuAlaGlnIleProProHisAlaValMet 628  
DB 1621 GCCGAGGTGTCCAGTTTGCAGCCTGCAACCTGCGCACAGATACCAACCCCGCGGTGATG 1680  
QY 629 ValArgProAspThrAsnIlePheThrValTyrGlyLeuLeuAspLysAlaGln 646  
DB 1681 GTCCGCCCCGACACCAACATCTTCCCGTGTATGACTGTGTCGACAGAGGCCGAG 1734

## RESULT 3

AY420524

## LOCUS

AY420524

## DEFINITION

Mus musculus MF12 gene, VIRTUAL TRANSCRIPT, partial sequence,

genomic survey sequence.

## ACCESSION

AY420524

## VERSION

AY420524.1

## KEYWORDS

GSS.





Db 1201 GAGTTTCGGGGCAAGCCCTCTGCGCATGCCGCTTTCGGCAGCCCTGAGGCTGGGATGTC 1260  
QY 489 ProValGlyAlaLeuLeuGlnArgGlyPheIleArgProLysAspCysAspValLeuThr 508  
Db 1261 CCCGTGGGGCCCTTATTTCAGAGAGGCTTCATCCGGCCCAAGGACTGTGAGCTCTCACA 1320  
QY 509 AlaValSerGluPhePheAsnAlaSerCysValProValAsnAsnProLysAsnTyrPro 528  
Db 1321 GNN 1380  
QY 529 SerSerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGlyAsn 548  
Db 1381 TCTCTGCTGTGTCAGTCTGCTGGGGAGCAGCAGCGCCGCAACAAAGTGTGGGCAAC 1440  
QY 549 SerGlnGluArgTyrTyrGlyTyrArgGlyValaPheArgCysLeuValGluAsnAlaGly 568  
Db 1441 AGCCAGAGCGGTATTACGGCTACCGCGCGCCTTCAGGTGCTCTGGTGAGATGCGGGT 1500  
QY 569 AspValAlaPheValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSerGlu 588  
Db 1501 GAGGTTCCTTCGTGAGGCACACACCGCTTTTGACAAACAAANNNNNNNNCAATTCCGAG 1560  
QY 589 ProTTPAlaAlaGluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAlaArg 608  
Db 1561 CCTTGGCTGCTGAGCTCAGGTGAGGAGCTATGAACCTGCTGTGCCCAACCGGCCCN 1620  
QY 609 AlaGluValSerGlnPheAlaAlaCysAsnLeuAlaGlnIleProProHisAlaValMet 628  
Db 1621 NCCGAGGTGCCAGTTTCAGCTGCACTGCACTGGNNNNNNNNNNNNNNNNNNNNNN 1680  
QY 629 ValArgProAspThrAsnIlePheThrValTyrGlyLeuLeuAspLysAlaGln 646  
Db 1681 NTCCGGCCCGACACCAACATCTTCACCGTGTATGAGTCTGTGCAAGGCCAG 1734

## RESULT 5

BU185213  
LOCUS AGENCOURT\_6910960 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:5952201  
5', mRNA sequence.

ACCESSION BU185213

VERSION BU185213.1 GI:22699197

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1023)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Place: LNCM2139 row: n column: 10

High quality sequence stop: 692.

Location/Qualifiers

1. .1023

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5952201"

/tissue type="ductal carcinoma, cell line"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH\_MGC\_110"

/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI;

Site 2: EcoRI; cDNA made by oligo-dT priming.

Directionally cloned into EcoRI/XhoI sites using the

following 5' adaptor: GGCACGAG(G). Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,02e-119 Length: 1023  
Score: 1245.00 Matches: 235  
Percent Similarity: 98.76% Conservative: 3  
Best Local Similarity: 97.51% Mismatches: 3  
Query Match: 31.63% Indels: 0  
DB: 13 Gaps: 0

US-10-049-957-4 (1-738) x BU185213 (1-1023)

QY 1 MetArgGlyProSerGlyAlaLeuTyrLeuLeuLeuAlaLeuArgThrValLeuGlyGly 20  
Db 90 ATGCGGGTCCGAGCGGGCTCTGTGGCTGCTCTGGCTCTGGCACCCTGCTGGTGGC 149  
QY 21 MetGluValArgTyrCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
Db 150 ATGGAGGTGCGGTGGTGGCCACCTCGACCCAGACGACCAAGTGCAGCAACATGAGC 209  
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
Db 210 GAGGTCTTCGGGAAGCGGGCATCCAGCCCTCTCTCTGTCGGGGCAGCTCCGCC 269  
QY 61 AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80  
Db 270 GACCACTGGCTCAGCTCATCGCGCCAGAGGCTGACGCCATCACTCTGGATGAGGA 329  
QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100  
Db 330 GCATCTATGAGCGGGGAAGAGCAGCGCCCTGAAAGCCGTGGTGGCGCAAGTGTACGAT 389  
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValAlaValArgSerHisValThr 120  
Db 390 CAAGAGTTCGTACCTCTCTTACCGCGTGGCTGTGGTCAGGAGGAGCTCCCATGTGACC 449  
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140  
Db 450 ATTGACACCTGAAAGCGGTGAAGTCTGCCACAGCGGCATCAATGCCACAGTGGGCTGG 509  
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
Db 510 AACGTGCCGCTGGGTACCTGTGGTGGAGAGCGGCCCTCTCGGTGATGGGCTGCGATGTA 569  
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyValaGlyGluThrSer 180  
Db 570 CTCAAAGCTGTACGACGACTATTTTGGGGGACGCTGGTCCCCGGGGCAGAGACCACT 629  
QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGlyGlyValCysAsp 200  
Db 630 TACTCTGAGTCCCTCTGTGCGCTCTGCGAGGGGTGACAGCTCTGGGAAGGGGTGTGTGAC 689  
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGly 220  
Db 690 AAGAGCCCCCTGGAGAGATCTACGACTACAGCGGGGCTTTTCGGTGGCTGGCGGAAGG 749  
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
Db 750 GCAGGGGAGCTGGCTTTTGTGAAGCAGACAGCAGTACTGGAGAACACAGGATGAAGATCC 809  
QY 241 Leu 241  
Db 810 ATC 812

## RESULT 6

BQ929045

LOCUS

DEFINITION

AGENCOURT\_8967315 NCI\_CGAP\_Mam2 Mus musculus cDNA clone

mRNA linear

937 bp

EST 20-AUG-2002

IMAGE:6484854 5', mRNA sequence.

ACCESSION  
BQ929045  
VERSION  
BQ929045.1  
KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS  
TITLE  
JOURNAL  
COMMENT  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLMW14025 row: 1 column: 07

High quality sequence stop: 670.

FEATURES

Location/Qualifiers  
1..937  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strains="FVB/N-3"  
/db\_xref="taxon:10090"  
/clone="IMAGE:6484854"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="5 months"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Mam2"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by Life Technologies. Investigator  
providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:

Pred. No.: 1,11e-119 Length: 937  
Score: 1244.00 Matches: 241  
Percent Similarity: 88.67% Conservative: 25  
Best Local Similarity: 80.33% Mismatches: 32  
Query Match: 31.61% Indels: 3  
DB: 13 Gaps: 1

US-10-049-957-4 (1-738) x BQ929045 (1-937)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20  
DB 28 ATGAGGCTCTGAGCGGTGACTTTTGGTGACTCTCTGCTCGGCACTGTCTGTGTG 87  
QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
DB 88 ATGGAGGTGTCAGTGTGTACCATCTCAGACCGAGAGCAGAGAGTGCAGACATGAGC 147  
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
DB 148 GAGGCTTCCAGGAGCTGGCATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 207  
QY 61 AspHisCysValGlnLeuIleAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80  
DB 208 GACCACTGTGTCAGTCTATCAGGAACAAAGAGAGAGATGCCATCCCTGTGATGGAGG 267  
QY 81 AlaIleTyrGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100  
DB 268 GCATCTATGAGCAGGAGGAGAGCAGCGCTTGAAGCAGTGTGTGGGAGTCTATGAC 327  
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValArgSerHisValThr 120  
DB 328 CAAGACATTGGGACTTCTATTATGCGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 387

QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTyr 140  
DB 388 ATCAACACCTTGAAGGGCGTCAAGTCTGCGCACACAGGCANTTACCGGACTGTGGGCTGG 447  
QY 141 AsnValProValGlyTyrLeuValLysSerGlyArgLeuSerValMetGlyCysAspVal 160  
DB 448 AACGTGCTGTCGTTACCTCTGAGAGAGCGCCATCTGTCTGATGATGGCTGTGATGTG 507  
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180  
DB 508 CTCAAAGCCGTTGGTGATTATTTTGGAGGCAGCTGTCTCTTGGAAACAGGAGAACCCAGC 567  
QY 181 TyrSerGluSerLeuLysCysArgLeuLysSerGlyArgGlySerSerGlyGluValCysAsp 200  
DB 568 CATTCCGAGTCCCTCTGTGCGCTCTGCGGTGGGCACTCTTCTGGGCACAATGTGTGTGAC 627  
QY 201 LysSerProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGly 220  
DB 628 AAGAGTCCCTTAGAGAGATACTACGACTACAGTGGAGCCTTCCGGTGGCTTGGCGAAGGA 687  
QY 221 AlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThr 240  
DB 688 GCCGTGACGTGGCTTCTGTAAGCACAGCACAGTGTCTGGAATAATCTGATGGAACACC 747  
QY 241 LeuProSerTrpGlyGlnAlaLeuLysSerGlnAspPheGluLeuLeuLysCysArgAspGly 260  
DB 748 CTGCTCTCTCTGGGCAAGTCTCTGATGTCAGAGACTTCCAGGACTATGATGGAAGATGCG 807  
QY 261 SerArgAlaAspValThr-GluTrpArgGlnCysHisLeuAlaArgValProAlaHisAl 280  
DB 808 AGCCGAGCGGACATCACTTTGANGGGAGAGCTTGCACCTGGCCCAAGGGGCTGCTCATGC 867  
QY 280 aValValValArgAlaAspThr---AspGlyGlyLeuIlePheArgLeuLeuAsnGlu 298  
DB 868 TGGGGGGGGCGGGGGG-ACATGGAATGGGGGCTCTCTATTCCACTGGCTTAACGAA 924

RESULT 7

BUI68950

LOCUS

DEFINITION

AGENCOURT\_8076230 NIH\_MGC\_110 Homo sapiens

5', mRNA sequence.

ACCESSION

BUI68950

VERSION

BUI68950.1

GI:22682934

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 998)

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LUCM2317 row: 9 column: 12

High quality sequence stop: 626.

Location/Qualifiers

1..998

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6084899"

/tissue\_type="ductal carcinoma, cell line"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_110"

FEATURES

source

/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
Pred. No.: 6,79e-119 Length: 998  
Score: 1237.00 Matches: 237  
Percent Similarity: 96.36% Conservative: 1  
Best Local Similarity: 95.95% Mismatches: 6  
Query Match: 31.43% Indels: 3  
DB: 13 Gaps: 1

US-10-049-957-4 (1-738) x BU168950 (1-998)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20  
Db 90 ATGGGGGTCGAGCGGGCTCTGTGGCTCTCTGGCTCTGGCAGCGGCTCGGTGGC 149  
QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
Db 150 ATGGAGGTGGGTGGTGGCGCACCTCGGACCCAGAGCAGCACAAAGTGGCGCAATGAGC 209  
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
Db 210 GAGGTCTTCGGGAAGCGGCATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 269  
QY 61 AspHisCysValGlnLeuIleAlaAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80  
Db 270 GACCACTGCTCAGCTCATCGGGCCAGAGGCTGACGCCATCATCTCGGATGAGGA 329  
QY 81 AlaIleTrpGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100  
Db 330 GCCATCTATGAGCGGGAAGGAGCAGCGCTGAAGCGGTGGTGGCGAAGTGTACGAT 389  
QY 101 GlnGluValGlyThrSerTyrTyrAlaValAlaValAlaValArgSerHisValThr 120  
Db 390 CAAGAGTCCGTACCTCTATTACCGCGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 449  
QY 121 IleAspThrLeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrp 140  
Db 450 ATTGACACCTCGAAGCGGTGAAGTCTGTGCCACACGGGCTATCATGTGCACAGTGGCTGG 509  
QY 141 AsnValProValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspVal 160  
Db 510 AACGTGCGCGTGGCTACCTGTGTGGAGAGCGCGCGCTCTCGGTGTGTGTGTGTGTGT 569  
QY 161 LeuLysAlaValSerAspTyrPheGlyGlySerCysValProGlyAlaGlyGluThrSer 180  
Db 570 CTAAAGCTGTACGACGATATTTTGGGGGCGAGTGGCTGCCGGGGCGAGAGACGAGT 629  
QY 181 TyrSerGluSerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAsp 200  
Db 630 TACTCTGAGTCCCTCTGTCTGCTCTGTGAGGGGTGACAGCTCTGGGAAGGGGTGTGTGAC 689  
QY 201 LysSer-ProLeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGlu 220  
Db 690 AAGAGCCCCCTGGAGAGATACGACTACAGCGGGCCCTTCGGGTGGCTGGCGGAAGG 749  
QY 220 yAlaGlyAspValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLys 240  
Db 750 GGCAGGGGAGCTGGCTTTTGTGAAGCAGACCCCGGTACTGTGAGAACACGAGATGAATCC 809  
QY 240 rLeuPro-----SerTrp 244  
Db 810 ATCCCCAAGGAACATGG 828

RESULT 8

## BE274850

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

BE274850 819 bp mRNA linear EST 13-JUL-2000  
601122847F1 NIH\_MGC\_20 Homo sapiens cDNA clone IMAGE:3347123 5',  
mRNA sequence.  
BE274850 BE274850.1 GI:9149797  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTF  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LLCW137 row: m column: 12  
High quality sequence stop: 722.

Location/Qualifiers  
1..819  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3347123"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_20"  
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dt priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.1e-117 Length: 819  
Score: 1224.00 Matches: 240  
Percent Similarity: 94.16% Conservative: 2  
Best Local Similarity: 93.39% Mismatches: 8  
Query Match: 31.10% Indels: 7  
DB: 10 Gaps: 2

US-10-049-957-4 (1-738) x BE274850 (1-819)

QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGlyGly 20  
Db 50 ATGGGGGTCGAGCGGGCTCTGTGGCTCTCTGGCTCTGGCAGCGGCTCGGTGGC 109  
QY 21 MetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCysGlyAsnMetSer 40  
Db 110 ATGGAGGTGGGTGGTGGCGCACCTCGGACCCAGAGCAGCACAAAGTGGCGCAATGAGC 169  
QY 41 GluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCysValArgGlyThrSerAla 60  
Db 170 GAGGCTTCCGGGAAGCGGGCATCCAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 229  
QY 61 AspHisCysValGlnLeuIleAlaAlaGlnGluAlaAspAlaIleThrLeuAspGlyGly 80  
Db 230 GACCACTGCTCAGCTCATCGGGCCAGAGGCTGACGCCATCATCTCGGATGAGGA 289  
QY 81 AlaIleTrpGluAlaGlyLysGluHisGlyLeuLysProValValGlyGluValTyrAsp 100  
Db 290 GCCATCTATGAGCGGGAAGGAGCAGCGGCTTGAAGCGGTGGTGGCGAAGTGTACGAT 349









Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.W., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodrigues, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schneringer, A., Schein, J.E., Jones, S.J., and Marra, M.A.

**TITLE**  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

**PUBMED**  
12477932

**REFERENCE**  
2 (bases 1 to 2314)

**AUTHORS**  
Strausberg, R.

**TITLE**  
Direct Submission

**JOURNAL**  
Submitted (20-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

**REMARK**  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

**COMMENT**  
Contact: MGC help desk

Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 40 Row: h Column: 16  
This clone has the following problem: no 5' EST match.

**FEATURES**  
Location/Qualifiers

1..2314  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3968762"  
/tissue\_type="Mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."  
/clone\_lib="NCI CGAP\_Mam1"  
/lab\_host="DH105"  
/note="Vector: pCMV-SPORT6"

# ORIGIN

## Alignment Scores:

Pred. No.: 1,59E-114 Length: 2314  
Score: 1202.00 Matches: 301  
Percent Similarity: 53.62% Conservative: 114  
Best Local Similarity: 38.89% Mismatches: 265  
Query Match: 30.54% Indels: 94  
DB: 11 Gaps: 32

US-10-049-957-4 (1-738) x BC020295 (1-2314)

Qy 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGly--- 19  
|||||

Db 40 ATGAGGCTCACCGTGGGTGGCCCTGCTGGCCCTGGCGTGGCCCTGGCGCTGTGTGGCTGTC 99  
Qy 20 ---GlyMetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCys----- 36  
Db 100 CTGACAAAACGGTCAAATGTGGCAGTGTGCAGACGACGAGAAATACCAATGCATCAGC 159  
Qy 37 -----GlyAsnMetSerGluAlaPheArgGluAlaGlyIleGlnProSerLeuLeuCys 54  
Db 160 TTCGTGACACATGAGACCGCTCTTCCTGGCTGTATGGC-----CCCCGGCTTGGCTGT 213  
Qy 55 ValArgGlyThrSerAlaAspHisCysValGlnLeuIleAlaGlnGlnAlaAspAla 74  
Db 214 GTGAAGAAAACCTCTATCCGATTCATCAAGCCATTCTTCAAGTGAAGCCGATGCT 273  
Qy 75 IleThrLeuAspGlyGlyValaIleTrpGluAlaGly---LysGluHisGlyLeuLysPro 93  
Db 274 ATGACCTTGGATGGGGTGGGTGTACCTCCGCTGACTCCGAAACAACTGAAGGCC 333  
Qy 94 ValValGlyGluValTyrAsp-----GlnGluValGlyThrSerTyrTyrAlaValAla 111  
Db 334 GTGGCGGGAGTTTATGGATCAGTGGAACATCCACAGACCTACTACTACGCTGGCT 393  
Qy 112 ValValArgArgSerSerHisValThrIleAspThrLeuLysGlyValLysSerCysHis 131  
Db 394 GTGGTAAAGAGGGAACACAGACTTCAGCTGAACACAGCTCGAAGGCAAGAGTCTCGCCAC 453  
Qy 132 ThrGlyIleAsnArgThrValGlyTrpAsnValProValGlyTyrLeuValGluSerGly 151  
Db 454 ACAGCCCTGGGAAGGTCTGCAGGCTGGGTCAATCCCATTTGGCTTGTCTTTC-----TGT 507  
Qy 152 ArgLeuSerValMetGlyCysAspValLeuLysAlaValSerAspTyrPheGlyGlySer 171  
Db 508 AAGCTGTCGAGAGCCCGCAGTCTCTTGAGAAAGTGTGTCCAGTTTCTTCTCGGCGAGT 567  
Qy 172 CysValProGlyAlaGlyGluThrSerTyrSerGluSerLeuLeuCysArgGly 191  
Db 568 TGTGTCCCTCTGCAGATCCAGTGGCTTCCCCAAA---CTGTGTCAACTGTGCCAGGC 624  
Qy 192 AspSerSerGlyGluGlyValCysAspLysSerProLeuGluArgTyrTyrAspTyrSer 211  
Db 625 -----TGTGGCTGTCTTCCACTCAACCGCTTCTTGGCTTACGTA 663  
Qy 212 GlyAlaPheArgCysLeuAlaGlyAlaGlyAspValAlaPheValLysHisSerThr 231  
Db 664 GGCGCATTCAGTGTCTGAAGAGTGGCGTGGGATGTGGCTTTGTCAAGCACACACACC 723  
Qy 232 ValLeuGluAsnThrAspGlyLysThrLeuProSerTrpGlyGlnAlaLeuLeuSerGln 251  
Db 724 ATATTTGAG-----GTCTTGGCGGAGAGGCTGAC-----AGGGAC. 759  
Qy 252 AspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGluTrpArgGlnCys 271  
Db 760 CAATATGAACATGCTCTGCTTGAACAATACCCGCAAGCCAGTGGATCAGTATGAGGATTGC 819  
Qy 272 HisLeuAlaArgValProAlaHisAlaValValValArgAlaAspThr---AspGlyGly 290  
Db 820 TACCTGGCTCGGATCCCTCTCTCATGCTGTGTGGTCCGAAAAACAATGGCAAGGAGAC 879  
Qy 291 LeuIlePheArgLeuLeuAsnGluGlyGlnArgLeuPheSerHis---GluGlySerSer 309  
Db 880 TTGATCTGGAGATTCTCAAAAGTGCAACAGAAACACTTTTGGCAAGGCAATCAAAAGAC 939  
Qy 310 PheGlnMetPheSerSerGluAlaTyrGlyGlnLysAspLeuLeuPheLysAspSerThr 329  
Db 940 TTCCAACATGTTCAAGCTCTCTCTT-----GGGAAAGACCTGCTGTTTAAAGATTCTGCC 993  
Qy 330 SerGluLeuValProIleAlaThrGln---ThrTyrGluAlaTrpLeuGlyHisGluTyr 348  
Db 994 TTTGGCTGTTAAGGTTCCCCCAAGGATGACAGCTGACAGGCTGTACCTTGGCCATACTAT 1053  
Qy 349 LeuHisAlaMetLysGlyLeu-----LeuCysAspProAsnArgLeuPro---Pro 364  
Db 1054 GTCACCTGCCATTCCGAATCAGCAGAGGCGGTGTGCCCGGAGGCTCGATCGACAACCTCG 1113

QY 365 TyLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGlyAspMetalVal 384  
 Db : : : : :  
 1114 CCAGTGAAGTGTGTGCTGAGTCACTCGAGAGAACCAAGTGTGACGAGTGGAGCATC 1173  
 QY 385 AlaPheArgGlnArgLeuLysProGluIleGlnCysValSerAlaLysSerProGln 404  
 Db : : : : :  
 1174 ATC-----AGTGGGAAGATAGAGTGTGAGTCAGCAGACCACTGAG 1218  
 QY 405 HisCysMetGluArgIleGlnAlaGluGlnValAspAlaValThrLeuSerGlyGluAsp 424  
 Db : : : : :  
 1219 GACTGCATTGAAGATTGTGAACGAGAGCGACCGCATGACTTGTGGATGAGACAT 1278  
 QY 425 IleTyrThrAlaGlyLysIleGlyLeuValProAlaIleGlyGluHisTyrAlaPro 444  
 Db : : : : :  
 1279 GCCTACATTGCAGGC---CAGTGTGTCTAGTGCCTGTCTATGCGTCAGCAGTACTAC----- 1329  
 QY 445 GluAspSerSerAsn-----SerTyrTyr 452  
 Db : : : : :  
 1330 ---GAGAGCTCTAATTGTGCTCATCCCATCAACAAGGTATCTTCTCTAAAGGGTATTAT 1386  
 QY 453 ValValAlaValValArgAspSerSerHisAlaPheThrLeuAspGluLeuArgGly 472  
 Db : : : : :  
 1387 GCGTGTGCTGTGTGAAGCA---TCGGACATGATCATCCTGGACACACCTGAAGGC 1443  
 QY 473 LysArgSerCysHisAlaGlyPheGlySerProAlaGlyTrpAspValProValGlyAla 492  
 Db : : : : :  
 1444 AAGAAGTCTCTGCACACTGCGGTAGACAGAACCGTGTGTGGAACATCCCTATGGGCATG 1503  
 QY 493 LeuIleGlnArgGlyPheIleArgProLysAspCysAspValLeuThrAlaValSerGlu 512  
 Db : : : : :  
 1504 CTG-----TACACACGATCAACCACTGC-----AAATTCGATGAA 1539  
 QY 513 PhePheAsnAlaSerCysValProValAsnAsnProLysAsnTyrProSerSerLeuCys 532  
 Db : : : : :  
 1540 TTTTTCAGTCAAGCTGCGCTCCC---GGGTATGAGAAGAT-----TCCACCTCTGT 1590  
 QY 533 AlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGlyAsnSerGlnGluArg 552  
 Db : : : : :  
 1591 GACCTGTGTATTGGCCCA-----CTCAATGTCTCCGAACCAACAAGAGGAA 1638  
 QY 553 TyrTyrGlyTyrArgGlyAlaPheArgCysLeuValGluAsnAlaGlyAspValAlaPhe 572  
 Db : : : : :  
 1639 TATAATGGTTTACACAGGGGCTTTTCAGTGTCTCGTTGAGAAA---GGAGATGTAGCCTTT 1695  
 QY 573 ValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSerGluProTTPAlaAla 592  
 Db : : : : :  
 1696 GTGAACACCCAGACTGTCTGTGGATACACCGAAGGAAGAACCTCCCGAATGGGCTAAG 1755  
 QY 593 GluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAlaArgAlaGluValSer 612  
 Db : : : : :  
 1756 AATCTGAAGCAGGAAGACTTCGAGTTGCTCTGCTGCTGATGCGCACAGGAGCCTGTGAAA 1815  
 QY 613 GlnPheAlaAlaCysAsnLeuAlaGlnIleProProHisAlaValMetValArgProAsp 632  
 Db : : : : :  
 1816 GATTTTGCAGCTGCCACCTGGCCCAAGCTCCCAACCATCTGTGTCTCCACGAAGAAG 1875  
 QY 633 ThrAsnIlePheThrValTyrGlyLeuLeuAspLysAlaGlnAspLeuPheGlyAspAsp 652  
 Db : : : : :  
 1876 ---AAGCGACCCGGGTAAAGGCTGTACTGACTAGCCAGAGAGACTTTATTTTGGGGGAAGT 1932  
 QY 653 HisAsnLysAsnGlyPheLysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuLeu 672  
 Db : : : : :  
 1933 GACTGCACCGGCAATTCTGTTGTTCAGTCTACCTACACC-----AAGGACCTTCTG 1983  
 QY 673 PheLysAspAlaThrValArgAlaValProValIleGlyLysThrThrTyrArgGlyTyr 692  
 Db : : : : :  
 1984 TTCAGGATGACACCAAAATTTTCGTAAACTTCCAGAGGGGTACCAACCTGAAAAATAC 2043  
 QY 693 LeuGlyLeuAspTyrValAlaAlaLeuGluGlyMetSerSerGlnGlnCysSerGlyAla 712  
 Db : : : : :  
 2044 TTAGGACCGGAGTACATGCAATCTGTGGTAAACATG-----AGGAAGTGTCTCAACCTCA 2097

QY 713 AlaAla-----ProAlaProGlyAlaProLeu 721  
 Db CGAGCTCTGGAAGCCTGCACCTTTCCACAAACATTAATAATCCAGAGGTGGTGGCCACTG 2157  
 QY 722 ---LeuProLeuLeuLeuProAlaLeuAlaAlaArgLeuLeu 734  
 Db : : : : :  
 2158 TGTGTGAGACAGATGCTCTCTCCGTGGCCCATGGGCTTCTC 2199  
 RESULT 11  
 AK085754 2316 bp mRNA linear HTC 20-SEP-2003  
 LOCUS Mus musculus 10 days lactation, adult female mammary gland cDNA,  
 DEFINITION RIKEN full-length enriched library, clone:D730049G12  
 product:transferrin, full insert sequence.  
 ACCESSION AK085754  
 VERSION AK085754.1 GI:26351790  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 1 Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2  
 REFERENCE  
 1 Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 20499374  
 MEDLINE 20499374  
 PUBMED 11042159  
 3  
 REFERENCE  
 1 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kitsumai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, S.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)  
 20530913  
 MEDLINE 20530913  
 PUBMED 11076861  
 4  
 REFERENCE  
 1 The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 NATURE 409, 685-690 (2001)  
 5  
 REFERENCE  
 1 The FANTOM Consortium and the RIKEN Genome Exploration Research  
 Group Phase I & II Team.  
 TITLE Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 NATURE 420, 563-573 (2002)  
 6 (bases 1 to 2316)  
 ADACHI, J., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P.,  
 FUKUDA, S., FURUNO, M., HANAGAKI, T., HARA, A., HASHIZUME, W.,  
 HAYASHIDA, K., HAYAMOTO, N., HIRAMOTO, K., HIRAKAWA, T., HIROZANE, T.,  
 HORI, F., INOTANI, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T.,  
 KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KOUNDA, M.,  
 KOYA, S., KURIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M.,  
 NAKAMURA, M., NISHI, K., NUMAZAKI, R., NUMAZAKI, R., OHNO, M., OHSAITO, N.,  
 OKAZAKI, Y., SAITO, R., SAITO, H., SAKAI, C., SAKAI, K., SAKAZUME, N.,  
 SANO, H., SASAKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T.,  
 SUGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKASHIRA, S.,  
 TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A.,  
 MURAMATSU, M. and HAYASHIZAKI, Y.  
 Direct Submission



```
Db 1190 ATC-----ACTGAGGGAAGATAGATGAGTGTGAGTCAGCAGAGACCACTGAG 1234
Qy 405 HisCysMetGluArgIleGlnAlaGluInValAspAlaValThrLeuSerGlyGluAsp 424
   |||:|||||:|||||:
Db 1235 GACTGCTTGAAGAAGATTGTCAACGGAAGCGGACGCCATGATCTTTGGATGGAGACAT 1294
Qy 425 IleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGluHisTyrAlaPro 444
   |||:|||||:|||||:|||||:|||||:|||||:
Db 1295 GCTTACATTGCAGC---CAGTGTGGTCTAGTGCCTGTCATGGCAGAGTACTAC----- 1345
Qy 445 GluAspSerSerAsn-----SeryTyrTyr 452
   |||:|||||:
Db 1346 --GAGAGCTCTAATTGTGCGCATCCCATCACAAAGGTATCTTCTTAAAGGTATTAT 1402
Qy 453 ValValAlaValValArgAspSerSerHisAlaPheThrLeuAspGluLeuArgGly 472
   |||:|||||:|||||:|||||:|||||:|||||:
Db 1403 GCGGTGGCTGTGGTAGGCA---TCGGACACTAGCATCATCCCTGGAACACCTGAAAGGC 1459
Qy 473 LysArgSerCysHisAlaGlyPheGlySerProAlaGlyTyrPaspValProValGlyAla 492
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1460 AAGAAGTCCTGCCACACTGCGGGTAGACAGAACCGCTGGTTGGAAACATCCCTATGGGCATG 1519
Qy 493 LeuIleGlnArgGlyPheIleArgProLysAspCysAspValLeuThrAlaValSerGlu 512
   |||:|||||:|||||:|||||:
Db 1520 CTG-----TACACAGCATCAACCACTGC-----AAATTCGATGAA 1555
Qy 513 PhePheAsnAlaSerCysValProValAsnAsnProLysAsnTyrProSerSerLeuCys 532
   |||:|||||:|||||:|||||:
Db 1556 TTTTTCAGTCAAGCTGCGCTCC---GGGTATGAGAAGAAAT-----TCCACCTCTGT 1606
Qy 533 AlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGlyAsnSerGlnGluArg 552
   |||:|||||:|||||:|||||:
Db 1607 GACCTGTGTATTGGCCCA-----CTCAATGTGCTCCGACCAACAAGAGGAA 1654
Qy 553 TyrTyrGlyTyrArgGlyAlaPheArgCysLeuValGluAsnAlaGlyAspValAlaPhe 572
   |||:|||||:|||||:|||||:
Db 1655 TATAATGTTACACAGGGCTTTCAGGTGCTCTGTTGAGAA---GGAGATGAGCTTT 1711
Qy 573 ValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSerGluProTyrPalaAla 592
   |||:|||||:|||||:|||||:|||||:
Db 1712 GTGAAACACCAGACTGTCTGGATATACACCGAAGGAAGAACCTCCGGAATGGGCTAAG 1771
Qy 593 GluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAlaAlaGluValSer 612
   |||:|||||:|||||:|||||:|||||:
Db 1772 AATCTGAAGCAGGAAGACTTCTGAGTTGCTCTGCTCCCTGATGGCACCGAGCCTGTGAAA 1831
Qy 613 GlnPheAlaAlaCysAsnLeuAlaGlnIleProProHisAlaValMetValArgProAsp 632
   |||:|||||:|||||:|||||:|||||:
Db 1832 GATTTTGGCAGCTGCCACCTGGCCCAAGCTCCAAACCATGTTGTGTCTCACGAAAGAG 1891
Qy 633 ThrAsnIlePheThrValTyrGlyLeuLeuAspLysAlaGlnAspLeuPheGlyAspAsp 652
   |||:|||||:|||||:|||||:
Db 1892 --AAGCAGCCCGGGTTAAGGCTGTACTGACTAGCCAGGAGACTTTATTGGGGGAAGT 1948
Qy 653 HisAsnLysAsnGlyPheLysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuLeu 672
   |||:|||||:|||||:|||||:
Db 1949 GACTGCACCCGCGCAATTTCTTTGTTTCAAGTCCACCACC-----AAGGACCTTCTG 1999
Qy 673 PheLysAspAlaThrValArgAlaValProValGlyGlyLysThrThrTyrArgGlyTyr 692
   |||:|||||:|||||:|||||:|||||:
Db 2000 TTCAGGATGACACCAAAATGTTTCGTTAAACTTCCAGAGGGTACCAACCTGAAATAATAC 2059
Qy 693 LeuGlyLeuAspTyrValAlaAlaLeuGluGlyMetSerSerGlnGlnCysSerGlyAla 712
   |||:|||||:|||||:|||||:
Db 2060 TTAGGACGGAGTACATGCAATCTGTCGGTAAACATG-----AGGAAGTCTCAACCTCA 2113
Qy 713 AlaAla-----ProAlaProGlyAlaProLeu 721
   |||:|||||:
Db 2114 CGACTCTCGAAGCTGCACTTTTCCAAACATTAATAATCCAAAGAGTGGTGGCCACTG 2173
Qy 722 ---LeuProLeuLeuLeuProAlaLeuAlaAlaArgLeuLeu 734
   |||:|||||:|||||:
Db 2174 TGTGGAGACAGATGCTCCCTCCCGTGGGCCCATGGGCTTCTC 2215
```

## RESULT 12

BC058218

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

FEATURES

source

BC058218 2332 bp mRNA linear HTC 17-DEC-2003  
Mus musculus transferrin, mRNA (cdna clone IMAGE:5388952).

BC058218.1 GI:37194857

HTC.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2332)

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Donald, M.F., Casavant, T.L.,  
Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McSwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalus, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2332)

Strausberg, R.

Direct Submission

Submitted (15-SEP-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

NIH-MGC Project URL: <http://mgs.nci.nih.gov>

Contact: MGC help desk

Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth

Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)

Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,  
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dierich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,  
Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgeon, C., Vogt, J.D., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 127 Row: c Column: 23

This clone has the following problem: no 5' EST match.

Location/Qualifiers

1..2332

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="NMRI"

/db\_xref="taxon:10090"

```
/clone="IMAGE:5388952"  
/tissue_type="Mammary tumor. WAP-Tag model. 5 months old,  
gross tissue."  
/clone_lib="NCI_CGAP_Mam4"  
/lab_host="DH10B"  
/note="Vector: pCMV-SPORT6"
```

## ORIGIN

```
Alignment Scores:  
Pred. No.: 1.62e-114 Length: 2332  
Score: 1202.00 Matches: 301  
Percent Similarity: 53.62% Conservative: 114  
Best Local Similarity: 38.89% Mismatches: 265  
Query Match: 30.54% Indels: 94  
DB: 11 Gaps: 32  
  
US-10-049-957-4 (1-738) x BC058218 (1-2332)  
  
QY 1 MetArgGlyProSerGlyAlaLeuTrpLeuLeuAlaLeuArgThrValLeuGly--- 19  
DB 50 ATGAGGCTCACCGTGGGTGCCCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGTC 109  
  
QY 20 ---GlyMetGluValArgTrpCysAlaThrSerAspProGluGlnHisLysCys----- 36  
DB 110 CCTGACAAACCGTCAAAATGGTGGCAGTGTGAGAGCAGAGATACCAATGTCATCAGC 169  
  
QY 37 -----GlyAsnMetSerGluAlaPheArgGluAlaGlyLeGlnProSerLeuLeuCys 54  
DB 170 TTCGTGTGACCATGATGAAGACCGTCTCGCTGGCTGGCTGGCTGGCTGGCTGGCTG 223  
  
QY 55 ValArgGlyThrSerAlaAspHisCysValGlnLeuLeuAlaAlaGlnGluAlaAspAla 74  
DB 224 GTGAGAAACCTCTATCCGATGTCATCAGGCGCATTTCTGCAAGTGAAGCCGATGCT 283  
  
QY 75 IleThrLeuAspGlyCysAlaIleTrpGluAlaGly---LysGluHisGlyLeuLysPro 93  
DB 284 ATGACCTGGATGGGGTGGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 343  
  
QY 94 ValValGlyGluValTyrAsp-----GlnGluValGlyThrSerTyrTrpAlaValAla 111  
DB 344 GTGGCGGGGAGTATTTATGGATCAGTGGGAACATCCACAGACCTACTACTACGCTGGCT 403  
  
QY 112 ValValArgArgSerHisValThrIleAspThrLeuLysGlyValLysSerCysHis 131  
DB 404 GTGGTAAGAAGGACAGACTTCAGCTGACAGCTGCAAGGCAAGAGTCTCGCCAC 463  
  
QY 132 ThrGlyIleAsnArgThrValGlyTrpAsnValProValGlyTyrTrpValGluSerGly 151  
DB 464 ACAGGCTGGGAAGTCTGACAGGCTGGGTCTATCCCATTTGGCTTCTCTTC-----TGT 517  
  
QY 152 ArgLeuSerValMetGlyCysAspValLeuLysAlaValSerAspTyrPheGlyGlySer 171  
DB 518 AAGCTGTGGAGCCCGCAGTCTCTTGAGAAAGTGTGTCAGTATTTCTTCTCGGGCAGT 577  
  
QY 172 CysValProGlyAlaGlyGluThrSerTyrSerGluSerLeuCysArgLeuCysArgGly 191  
DB 578 TGTGTCCCTGTGCAAGTCCAGTGGCTTCCCAA---CTGTGCAACTGTGCCAGGC 634  
  
QY 192 AspSerSerGlyGluGlyValCysAspLysSerProLeuGluArgTyrTrpAspTyrSer 211  
DB 635 -----TGTGGCTGCTCTCCCATCAACCGTCTCTTTGGCTACGTA 673  
  
QY 212 GlyAlaPheArgCysLeuAlaGluGlyAlaGlyAspValAlaPheValLysHisSerThr 231  
DB 674 GCGCATTCAGTGTCTGAAGATGGCGTGGGGATGTGGCTTTTGTCAAGCAGCACCAACC 733  
  
QY 232 ValLeuGluAsnThrAspGlyLysThrLeuProSerTrpGlyGlnAlaLeuLeuSerGln 251  
DB 734 ATATTTGAG-----GTCCTTCCCGAGAGAGGCTGAC-----AGGGAC 769  
  
QY 252 AspPheGluLeuLeuCysArgAspGlySerArgAlaAspValThrGluTrpArgGlnCys 271  
DB 770 CAATATGAATCTCTCTGCTTGCACAAATACCCGCAAGCCAGTGTGATGAGGATTGC 829
```

```
QY 272 HisLeuAlaArgValProAlaHisAlaValValArgAlaAspThr---AspGlyGly 290  
DB 830 TACTGGCTCGATCCCTCTCTCATGCTGTGTGGCTCGAAAAACAATGGCAAGAACAC 889  
  
QY 291 LeuIlePheArgLeuLeuAsnGluGlyGlnArgLeuPheSerHis---GluGlySerSer 309  
DB 890 TTGATCTGGGAGATTCTCAAAAGTGGCAGGAAACACTTTTGGCAAGGCAATCAAAAGAC 949  
  
QY 310 PheGlnMetPheSerSerGluAlaTyrGlyGlnLysAspLeuLeuPheLysAspSerThr 329  
DB 950 TTCCAACATGTTTCAGCTCTCTCTT-----GGGAAAGACCTGCTGTTTAAAGATTCTGCC 1003  
  
QY 330 SerGluLeuValProIleAlaThrGln---ThrTyrGluAlaTyrLeuGlyHisGluTyr 348  
DB 1004 TTTGGCTGTTAAGGGTCCCCCAAGGATGAGTACAGGCTGTACTTGGCCATAACTAT 1063  
  
QY 349 LeuHisAlaMetLysGlyLeu-----LeuCysAspProAsnArgLeuPro---Pro 364  
DB 1064 GTCACCTGCATTCGGAATCAGCAGGAAGCGGTGTGCCGGAGGGCTCGATCGACAACCTCG 1123  
  
QY 365 TyrLeuArgTrpCysValLeuSerThrProGluIleGlnLysCysGlyAspMetAlaVal 384  
DB 1124 CCAGTGAAGTGTGTGACCTGAGTCACTGAGTGGAGAGAACCAAGTGTGACGAGTGGAGCATC 1183  
  
QY 385 AlaPheArgArgGlnArgLeuLysProGluIleGlnCysValSerAlaLysSerProGln 404  
DB 1184 ATC-----AGTGAAGGAAGATAGATGTGAGTGCAGCAGAGACCACTGAG 1228  
  
QY 405 HisCysMetGluArgIleGlnAlaGluValAlaValThrLeuSerGlyGluAsp 424  
DB 1229 GACTGCATTTGAAAGATTGTGAACGGAGACGGACGCCATGACTTTTGGATGGAGCAT 1288  
  
QY 425 IleTyrThrAlaGlyLysLysTyrGlyLeuValProAlaAlaGlyGluHisTyrAlaPro 444  
DB 1289 GCCTACATTTGAGGC---CAGTGTGTCTAGTGTCTGTGATGGCAGAGTACTAC----- 1339  
  
QY 445 GluAspSerSerAsn-----SerTyrTyr 452  
DB 1340 ---GAGAGCTCTAATTTGGCCATCCCATCAACAAGGTATCTTCTTAAAGGGTATTAT 1396  
  
QY 453 ValValAlaValArgArgAspSerHisAlaPheThrLeuAspGluLeuArgGly 472  
DB 1397 GCCGTGGCTGTGGTGAAGCA---TCGGACACTAGCATCCTCGAACAACCTGAAAGGC 1453  
  
QY 473 LysArgSerCysHisAlaGlyPheGlySerProAlaGlyTrpAspValProValGlyAla 492  
DB 1454 AAGAAGTCTCTGCCACACTGGGGTAGACAGAACCGTGTGTGGAAACATCCCTATGGGCATG 1513  
  
QY 493 LeuIleGlnArgGlyPheIleArgProLysAspCysAspValLeuThrAlaValSerGlu 512  
DB 1514 CTG-----TACAACAGGATCAACCACTGC-----AAATTCGATGAA 1549  
  
QY 513 PhePheAsnAlaSerCysValProValAsnAsnProLysAsnTyrProSerSerLeuCys 532  
DB 1550 TTTTTCAGTCAAGCTGGCTCCC---GGGTATGAGAAGAT-----TCCACCCCTCTGT 1600  
  
QY 533 AlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGlyAsnSerGlnGluArg 552  
DB 1601 GACCTGTGTATTGGCCCA-----CTCAAAATGTGTCTCCGAAACAACAAGAGGAA 1648  
  
QY 553 TyrTyrGlyTyrArgGlyAlaPheArgCysLeuValGluAlaGlyAspValAlaPhe 572  
DB 1649 TATAATGTTTACACAGGGGCTTTTCAAGTGTCTGTGTGAGAAA---GGAGATGTAGCCCTTT 1705  
  
QY 573 ValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSerGluProTrpAlaAla 592  
DB 1706 GTGAACACACAGACTGTCTCGATTAACACCGAAGGAAGAACCCCTGCCGATGGGCTAAG 1765  
  
QY 593 GluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAlaArgAlaGluValSer 612  
DB 1766 AATCTGAAGCAGGAAGACTTCAGTGTGTCTCTGCTGTGATGTCAGCAGGACCTGTGTGAA 1825
```





265 ValThrGluTrpArgGlnCysHisLeuAlaValProAlaHisAlaValValValArg 284  
 760 GTGGATCAGTATGAGGACTGTACCTAGCCCGGATCCCTTCATGCTGTGTGGCTCGA 819  
 285 ---AlaAspThrAspGlyLeuLeuPheArgLeuLeuAsnGluGlyGlnArgLeuPhe 303  
 820 AATGGAGATGGCAAGAGGACTGTATCTGGGAGATCTCTCAAAAGTGGCTCAGGAACACATT 879  
 304 SerHis---GluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGlnLysAsp 322  
 880 GGCAAGGCAATCAAAAGACTTCCAACTGTTCGGCTCTCTCTT-----GGGAAGAC 933  
 323 LeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGln---ThrTyrGlu 341  
 934 CTGCTGTTTAAGATTCTGCTTTGGCTGTTCAGGGTCCCAAGATGAGTACAGG 993  
 342 AlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeu-----LeuCysAsp 358  
 994 CTGTACCTCGGCACAGCTATGTCTACTGCTTCATTCGAATTCAGGAGCGGTGTGCCG 1053  
 359 ProAsnArgLeuProProTyr---LeuArgTrpCysValLeuSerThrProGluLeuGln 377  
 1054 GAGGCTCCATCGACGCGCGCAGTGAATGCTGTGCATCGATCAACCAAGAGAGACC 1113  
 378 LysCysGlyAspMetAlaValAlaPheArgArgGlnArgLeuLysProGluLeuGlnCys 397  
 1114 AAGTGTGATGAGGCTGAGCGTCAGCAATGGCGAG-----ATAGAGTGT 1158  
 398 ValSerAlaLysSerProGlnHisCysMetGluArgIleGlnAlaGluGlnValAspAla 417  
 1159 GAGTCAGCAGAGACTGAGGACTGCAATGACAAGATTGTGAATGGAGAACAGATGCC 1218  
 418 ValThrLeuSerGlyCysGluAspIleTyrThrAlaGlyLysLysTyrGlyLeuValProAla 437  
 1219 ATGAGCTGGATGGAGGTTCCTGCATCATACATAGAGCG---CAGTGTGGATAGTGCCTC 1275  
 438 AlaGlyGluHisTyrAlaProGluAspSerSerAsn----- 449  
 1276 ATGGCAGAGATGATGATATCTTCGTGTACAAACCCCAATCAGATGCTTTCTCTAAA 1335  
 450 SerTyrTyrValValAlaValArgArgAspSerSerHisAlaPheThrLeuAspGlu 469  
 1336 GGTATTATCGCTGGCTGTGTGAAGGATCAGACTCC---AGCATCAACTGGAAACAAC 1392  
 470 LeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAlaGlyTrpAspValPro 489  
 1393 CTGAAAGGCAAGAAGTCTCTGCCATCTACTGGAGTAGACAGAACCGCGCTGGAACATCCCT 1452  
 490 ValGlyAlaLeuIleGlnArgGlyPheIleArgProLysAspCysAspValIleuThrAla 509  
 1453 ATGGCGCTGCTG-----TTACGAGATCAACCACTGC-----AAG 1488  
 510 ValSerGluPhePheAsnAlaSerCysValProValAsnAsnProLysAsnTyrProSer 529  
 1489 TTCGATGAATTTTTCAGTCAAGCTGTGCTCT---GGTATAAGAGAT-----TCC 1539  
 530 SerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLysCysValGlyAsnSer 549  
 1540 ACCCTCTGTGACCTGTATTGGCCGACCA-----AAATGTGCTCCGGAACAAC 1587  
 550 GlnGluArgTyrTyrGlyTyrArgGlyAlaPheArgCysLeuValGluAsnAlaGlyAsp 569  
 1588 AGAGAGGAGATATATGGTTATACAGGGCTTTCCAGTGCCTCGTTGAGAAG---GGAGAC 1644  
 570 ValAlaPheValArgHisThrThrValPheAspAsnThrAsnGlyHisAsnSerGluPro 589  
 1645 GTAGCTTTGTGAGCACCAGACTGTCTTGGAAACACAGGAGGAAGAACACTCTCTGCA 1704  
 590 TrpAlaAlaGluLeuArgSerGluAspTyrGluLeuLeuCysProAsnGlyAlaArgAla 609  
 1705 TGGGCTAAGATCTGAAGCAGGAGAGACTTCCAGCTGCTGTGCCCTGTGATGAGTACCAAGAAG 1764

610 GluValSerGlnPheAlaAlaCysAsnLeuAlaGlnIleProHisAlaValMetVal 629  
 1765 CTGTAAACCGATTGCCCACTGCACCTGCCCAAGCTCCAAACCATGTTGTGTGCTCA 1824  
 630 ArgProAsp-----ThrAsnIlePheThrValTyrGlyLeuLeuAspLysAlaGlnAsp 647  
 1825 CGAAAGAGAGGAGCGCCGGTTAGCACTGTG-----CTGACTGCCCAAGAGAT 1875  
 648 LeuPhe-----GlyAspAspHisAsnLysAsnGlyPheLysMetPheAspSerSerAsn 665  
 1876 TTATTTTGGAAAGGTGACAGGACTGCACTGGCAATTTCTGTGTTGTCGGGTCTTCCACC 1935  
 666 TyrHisGlyGlnAspLeuLeuPheLysAspAlaThrValAlaGlnAlaValProValGlyGlu 685  
 1936 -----AAGACCTTCTGTTACAGATGACACCAAGTGTTCACATAACTTCCAGAA 1986  
 686 LysThrThrTyrArgGlyTrpLeuGlyLeuAspTyrValAlaAlaLeuGluGlyMetSer 705  
 1987 GGTACCATCATATGAAGAGTACTTAGGACGAGTACTTGAAGCTGTGGAACATA--- 2043  
 706 SerGlnGlnCysSer----- 710  
 2044 ---AGGAAGTGTTCAACTCAGACTCCTAGAGCTGCACCTTTCCACAAAGTTAAAT 2100  
 711 -----GlyAlaAlaProAlaProGlyAlaProLeuLeuProLeuLeuPro 727  
 2101 CCAAGAGTGGTGCCACTGTGTGGAGGAGATGCCCCCGTGATCCATGGCTTCTCT 2160  
 728 AlaLeuAlaAla 731  
 2161 GGCCTCCATGCC 2172  
 RESULT 14  
 BC008530  
 LOCUS  
 DEFINITION Mus musculus lactotransferrin, mRNA (CDNA clone IMAGE:3484893).  
 ACCESSION BC008530  
 VERSION BC008530.1 GI:14250215  
 KEYWORDS HTC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2296)  
 Strausberg, R.B., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
 Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,  
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,  
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
 Butlerfield, Y.S., Krzywicki, M.I., Skalska, U., Smalios, D.E.,  
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 2296)  
 Strausberg, R.  
 Direct Submission  
 Submitted (25-MAY-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA



REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgaps-i@mail.nih.gov](mailto:cgaps-i@mail.nih.gov)  
 Tissue Procurement: Iohar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
 R. M.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 7 Row: c Column: 13  
 This clone has the following problem: no 5' EST match.

FEATURES  
source

Location/Qualifiers  
 1. 2296  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3484893"  
 /tissue\_type="Mammary tumor. WAP-TGF alpha model. 7 months  
 old, gross tissue."  
 /clone\_lib="NCI CGAP\_Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,88e-114 Length: 2296  
 Score: 1199.50 Matches: 292  
 Percent Similarity: 54.57% Conservative: 114  
 Best Local Similarity: 39.25% Mismatches: 253  
 Query Match: 30.48% Indels: 85  
 DB: 11 Gaps: 27

US-10-049-957-4 (1-738) x BC008530 (1-2296)

QY 7 AlaLeuTrpLeuLeuLeuAlaLeuArgThrValLeuGlyGlyMetGluValArgTrpCys 26  
 Db 52 GCCTTTGGACTGTCTAGCTAAGGCAACAACCT-----GTTTCATGCTGT 96  
 QY 27 AlaThrSerAspProGluGlnHisLysCysGlyAsnMetSerGluAlaPheArgGluAla 46  
 Db 97 GCTGTGTCAAAATCTGTAGGAAGAAATAATGTTAAGGTGGCAGAACGAGATGAGAAAGTG 156  
 QY 47 GlyIleGlnProSerLeuLeuCysValArgGlyThrSerAlaAspHisCysValGlnLeu 66  
 Db 157 GGT---GGCCCGCGCTCACTTGTGTCAAGAAATCTCCACCCGCGCAGTGCATCCAGGCC 213  
 QY 67 IleAlaAlaGlnGluAlaAspAlaIleThrLeuAspGlyGlyAlaIleThrGluAlaGly 86  
 Db 214 ATTGTGACAAACAGAGCTGATGTCATGACTCTTGTATGTTGATGTCGACTATGTCGATGCAGGA 273  
 QY 87 Lys---GluHisGlyLeuLysProValValGlyGluValTy-----AspGlnGluVal 103  
 Db 274 AAGCCCCCTTACAAACTGGCACTGTGGCAGCTGGAAGTCTACGGGACCAAGAGCAGCCC 333  
 QY 104 GlyThrSerTyTrpAlaValAlaValValArgArgSerHisValThrIleAspThr 123  
 Db 334 CGGACTCACTACTATGCGGTAGCAGTGTGAAGNACAGCAGTAACCTTTCATCTGAACCAA 393  
 QY 124 LeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrpAsnValPro 143  
 Db 394 CTCGAAGCGCTGAGTCTCTGCCACACCGCATTTGGCAGGAGTGGCGGTGGAAATCCCT 453  
 QY 144 ValGlyTyTrpLeuValGluSerGlyArgLeuSerValMetGlyCysAspValLeuLysAla 163  
 Db 454 ATAGGGACACTTCGTCCATACCTCAATTTGAATTTGGGCGCACCTGCATCCCTTTGAGGAAGCG 513  
 QY 164 ValSerAspTyTrpPheGlyGlySerCysValProGlyAlaGlyGluThrSerTySerGlu 183  
 Db 514 GTATCCAAAGTTCTTCTCAAGAGCTGTGTTCGGTGCCTCAAAAGATAGATTC---CCC 570  
 QY 184 SerLeuCysArgLeuLeuCysArgGlyAspSerSerGlyGlyGlyValCysAspLysSerPro 203  
 Db 571 AACCTGTGTAGTGTGTGTGCGAGG-----ACAGGAGCCCAAAATGTGCTTCTTCCCCC 624  
 QY 204 LeuGluArgTyTrpAspTyTrpSerGlyAlaPheArgCysLeuAlaGlyAlaGlyAsp 223  
 Db 625 GAGGAGCCATCATCTAGGTTATCTGGAGCTTGTGAGGTGTCTGAGAGACAATGCTGGAGAT 684  
 QY 224 ValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuProSer 243  
 Db 685 GTGGCTTTTACCAGAGGAAGCAGCGTATTATTAGGAG-----TTATCCAAAT 729  
 QY 244 TrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGlySerArgAla 263  
 Db 730 AAAGCCGAA-----AGGAGCCAGTACAGCTGCTCTGCCAGACAACCTCGAAG 780  
 QY 264 AspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAlaValValVal 283  
 Db 781 CCGGTGACAGATAACAGAGTGCACCTGGCCCAAGTCCCTTCACATGCTGCTGATCC 840  
 QY 284 ArgAlaAspThrAsp---GlyGlyLeuIlePheArgLeuLeuAsnGluGlyGlnArgLeu 302  
 Db 841 CGAAGCACGAATGACAAAGAGAGCCATCTGGAGCTTCTCGCCAGTCACAGGAGAAG 900  
 QY 303 Phe---SerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyTrpGlyGlnLys 321  
 Db 901 TTTGGAAAAAACAAGCATCGGGATTCCAGCTTTTGTCTCCCTCCCTCCG---GGACAGAAG 957  
 QY 322 AspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrProIleAlaThr 340  
 Db 958 GACCTGCTGTCAAGAGTCTGCCATTCCTTGTGAGGGTTCCTCCAGAGGTAGATGTA 1017  
 QY 341 GluAlaTrpLeuGlyHisGluTyTrpLeuHisAlaMetLysGlyLeu----- 355  
 Db 1018 GGGCTCTACCTGACCTTCAGCTTACACCATACACAGACCTGAATAAAGACGACGAG 1077  
 QY 356 -----LeuCysAspProAsnArgLeuProProTyTrpLeuArgTrpCysValLeuSerThr 373  
 Db 1078 GATGTGATAGCTCAAGGCCCGG-----GTCACTATGGTGTCCGTGGGCGAGT 1125  
 QY 374 ProGluIleGlnLysCysGlyAspMetAlaValAlaPheArgGlnArgGlnArgLeuLysPro 393  
 Db 1126 GAGGAGAACCGCAAGTGTGATCAGTGGAAACAGACGACGAGGCGAGG----- 1173  
 QY 394 GluIleGlnCysValSerAlaLysSerProGlnHisCysMetGluArgIleGlnAlaGlu 413  
 Db 1174 ---GTCACCTGCATCTCAATTCCTCCACCGGAAGACTGCATGTGCAATCATGAAGGGA 1230  
 QY 414 GlnValAspAlaValThrLeuSerGlyGluAspIleTyThrAlaGlyLysTyTrpGly 433  
 Db 1231 GATGCTGATGCGCATGAGCTGTGAGGCTATATCTACACTGCGGCG---AAGTCGGT 1287  
 QY 434 LeuValProAlaAlaGlyGluHisTyTrpAlaProGluAspSerSer----- 448  
 Db 1288 TTAGTTTCCAGTCTTGGCAGAGAACCCAGAAATCTCCAAAGCAATGGCTTGGATTGTGTG 1347  
 QY 449 -----AsnSerTyTrpValValAlaValAlaValArgAspSerSerHisAla 464  
 Db 1348 AACAGACCAAGTGAAGGGTACCTTGTGTAGCAGCAGGTTAGAGAAGAGATGCT---GGC 1404  
 QY 465 PheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAla 484  
 Db 1405 TTCACCTGAGGCTCTTTGAGAGGCAAGAAATCTCTGCCACACTGCCCTGGACAGGACCGCA 1464  
 QY 485 GlyTrpAspValProValGlyAlaIleGlnArgGlyPheIleArgProLysAspCys 504  
 Db 1465 GGTGGGAATCCCTCCATGGGCTGCTTGTGTAC-----CAGACCATGATCTCTGC 1512  
 QY 505 AspValLeuThrAlaValSerGluPhePheAsnAlaSerCysValProValAsnAsnPro 524

```
Db      1513  -----AAATTTAATGAGTCTTTAGCCAAAGCTGTGCCCTGGTCTGACCCC 1560
QY      525  LysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLys 544
Db      1561  -----TCCAATCTCTGTGCCCTGTGTATTGGTGTATGAGAGGGGTGAGAACAG 1611
QY      545  CysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyValAlaPheArgCysLeuVal 564
Db      1612  TGTGCTCCACAGCAAGAGAGATACCAAGCTACCTGGGGCTTTTAAGGTGTCTGGCT 1671
QY      565  GluAsnAlaGlyAspValAlaPheValArgHisThrValPheAspAsnThrAsnGly 584
Db      1672  GAGAAGGCAGGAATGTTGCAATTTTGAAGGACTCCACTGTCTTCGAGAATACTGACGG 1731
QY      585  HisAsnSerGluProTyrAlaGluLeuArgSerGluAspTyrGluLeuLeuCysPro 604
Db      1732  AAGAACAACCTGAAGAGTGGCTAGGAACCTTAAAGCTGAAGGACTTTTGAGCTTTT 1791
QY      605  AsnGlyAlaArgAlaGluValSerGlnPheAlaAlaCysAsnLeuAlaGlnIlePro 624
Db      1792  GATGACACCCGAAACCTGTGACTGAGGCTAGAACTGCCACTAGCCATAGCCCCAAC 1851
QY      625  HisAlaValMetValArgProAspThrAsnIlePheThrValTyrGlyLeuLeuAspLys 644
Db      1852  CATGCTGTAGTCTCTCGGACAGACAAAGGTGGAAGTCCCTTCACGAGGTGCTGCTG 1911
QY      645  AlaGlnAspLeuPheGlyAspAspHisAsnLysAsnGly-----Phe 658
Db      1912  ---CAGGTTTCAGTTTTGGG-----AGAAATGGACAGAGGTGTCAGGAGAGTTT 1956
QY      659  LysMetPheAspSerSerAsnTyrHisGlyGlnAspLeuPheLysAspAlaThrVal 678
Db      1957  TCCTGTTTCAGTCTAAAC-----AAAAACCTTCTGTTCAATGACAACTAGAG 2007
QY      679  ArgAlaValProValGlyGlyLysThrTyrArgGlyTrpLeuGlyLeuAspTyrVal 698
Db      2008  TGTCTGCCCAAGATCCCGCAACACCATCGGAGAGTACTCGGAAAGGAGTACGTC 2067
QY      699  AlaAlaLeuGluGlyMetSerSerGlnGlnCysSerGlyAlaAlaProAlaProGly 718
Db      2068  ATAGGACCGGACGGCTG-----AAGCAGTCTCC-----AGC 2100
QY      719  AlaProLeuLeu 722
Db      2101  TCCCACTCTCTG 2112

RESULT 15
ACCESSION AK036491
LOCUS     Mus musculus adult male bone cDNA, Riken full-length enriched
DEFINITION library, clone:9830118D19 product:lactotransferrin, full insert
sequence.
ACCESSION AK036491
VERSION   AK036491.1 GI:26331439
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1
AUTHORS   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE     Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL   Normalization and subtraction of cap-trapper-selected cDNAs to
MEDLINE   prepare full-length cDNA libraries for rapid discovery of new genes
PUBMED    Genome Res. 10 (10), 1617-1630 (2000)
REFERENCE 2
AUTHORS   Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE     Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
JOURNAL   Normalization and subtraction of cap-trapper-selected cDNAs to
MEDLINE   prepare full-length cDNA libraries for rapid discovery of new genes
PUBMED    Genome Res. 10 (10), 1617-1630 (2000)
```

11042159

PUBMED  
REFERENCE  
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Saaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL  
MEDLINE  
PUBMED

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

PUBMED  
REFERENCE  
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

6 (bases 1 to 2744)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, C., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

JOURNAL  
REFERENCE  
AUTHORS

7 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

8 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

9 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

10 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

11 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

12 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

13 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

14 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

15 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

16 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

17 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

18 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

19 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

20 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

21 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

22 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

23 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

24 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

25 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

26 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

27 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

28 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

29 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

30 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

31 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

32 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

33 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

34 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

35 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

36 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

37 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

38 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

39 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

40 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

JOURNAL  
REFERENCE  
AUTHORS

41 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

42 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

43 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

44 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

45 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

46 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

47 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

48 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

49 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

50 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

51 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

52 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

53 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

54 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

55 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

56 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

57 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

58 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

59 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

60 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

61 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

62 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

63 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

64 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

65 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

66 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

67 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

68 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

69 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

70 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

71 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

72 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

73 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

74 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

75 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

76 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

77 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

78 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

79 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

80 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

81 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

82 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

83 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

84 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

85 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

86 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

87 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 6

/protein\_id="BAC29450.1"  
 /db\_xref="GI:26331440"  
 /translation="MRLLIPSLIFLEALGLCLAKATTVMQCAVSNSEBEKCLRWNEM  
 RKVGPPSLSCVKKSQRCIQAITVNRADAMTLDGTFDAGPKPYKLRVAAEYVGT  
 KEQPRHYAVAVVKNSSNFHLNQLGRSCTHIGRSAGWKIPICTLRPLYNWNGPP  
 ASILEAVSKFFSKSCVGAQKQFPNLCSSCAGTANKCASSPEEPYSGAGLRCLR  
 ANAGDVAFTRGSTVFEELPNKARDQYKLLCPDNTWKPTVEYKECHLAQVPSHAVVDV  
 STNDKEALWELLROSOBKQKQASGFOLFASPSGOKDLLFKESAIIGRVPVQKVDV  
 GYLTYFTYTSITQNLNKQDDVIASKARYTMCVAGSEERKCDOWNRAGRVTICISF  
 PITTEDCIVAIMKGDADMSLDGYIYTAGCKGLVPVLAENQKSSKSGLDVCNRPVEG  
 YLAAVVRREDAGFTWSLRGKKSCHTAVDRTAGWNI PMGLLANOTRSCKPFNEFSQS  
 CAPGAPKSNLCALCIGDEKGENKCA PNSKERVQGYTGALRCLAEKAGNVAFLKSTV  
 LQNTDKNTEEWARNILKLDPELLCLDDTRKPVTEANCHLAIPAHVAVSRIDKVEV  
 LQVLVDQVQFGRNQRCPGEFCLFQSKTKNLLFNFDNTECLAKIPGKTTSEKYLKGE  
 2720..2725  
 /note="putative"  
 2744  
 /note="putative"  
 2744  
 /note="putative"

polya\_signal

polya\_site

## ORIGIN

## Alignment Scores:

Pred. No.: 1,06e-113 Length: 2744  
 Score: 1195.50 Matches: 291  
 Percent Similarity: 54.57% Conservative: 115  
 Best Local Similarity: 39.11% Mismatches: 253  
 Query Match: 30.37% Indels: 85  
 DB: 11 Gaps: 27

US-10-049-957-4 (1-738) x AK036491 (1-2744)

QY 7 AlaLeuTrrpLeuLeuAlaLeuArgThrValLeuGlyMetGluValArgTrpCys 26  
 Db 72 GCCCTTGAGCTCTGTCTAGCTAAGCAACAACCT-----GTTCAATGCTGT 116  
 QY 27 AlaThrSerAspProGluGlnHisLysCysGlyAsnMetSerGluAlaPheArgGluAla 46  
 Db 117 GCTGTCTCAAAATCTCAGGAAGAAAATGTTTAAGGTGGCAGAACGAGATGAAAGATG 176  
 QY 47 GlyIleGlnProSerLeuLeuCysValArgGlyThrSerAlaAspHisCysValGlnLeu 66  
 Db 177 GGT--GGCCCGCGCTCAGTGTGTCAAGAAATCTCCACCCGCGAGTCATCCAGGCC 233  
 QY 67 IleAlaAlaGlnGluAlaAspAlaIleThrLeuAspGlyGlyAlaIleTyrGluAlaGly 86  
 Db 234 ATTGTGACAAACAGAGCTGATGCTGCTCTTGTGATGTGGCACTATGTTGATGCAGGA 293  
 QY 87 Lys--GluHisGlyLeuLysProValValGlyGluValTyr-----AspGlnGluVal 103  
 Db 294 AAGCCCCCTACAAACTGCGACCTGTGGCAGCTGGAAGTCTACGGGACCAAGAGCAGCCC 353  
 QY 104 GlyThrSerTyrTyrAlaValAlaValArgArgSerSerHisValThrIleAspThr 123  
 Db 354 CGGACTCACTACTATCGGTGAGCAGTCTGTGAAGAACAGCAGTAACCTTCATCTGAACCA 413  
 QY 124 LeuLysGlyValLysSerCysHisThrGlyIleAsnArgThrValGlyTrpAsnValPro 143  
 Db 414 CTCGAAGGCTAGGTCTCTGCCACACCGCATTTGGCAGAGTGGCGGGTGGAAAATCCCT 473  
 QY 144 ValGlyTyrLeuValGluSerGlyArgLeuSerValMetGlyCysAspValLeuLysAla 163  
 Db 474 ATAGGACACTTCGTCCATCTGAATGGAATGGGCCACCTGCAFTCCCTTGAGGAAGCG 533  
 QY 164 ValSerAspTyrPheGlyGlySerCysValProGlyValaGlyGluThrSerTyrSerGlu 183  
 Db 534 GTATCAAGTCTCTTCAAGAGCTGTGTCCGGTGCACCAAGATAGATTC---CCC 590  
 QY 184 SerLeuCysArgLeuCysArgGlyAspSerSerGlyGluGlyValCysAspLysSerPro 203  
 Db 591 AACCTGTGTAGTCTGTGTGCAGG-----ACAGGAGCCAAACAATGTGCCTCTTCCCC 644  
 QY 204 LeuGluArgTyrTyrAspTyrSerGlyAlaPheArgCysLeuAlaGluGlyAlaGlyAsp 223  
 Db 1632 TGTGTCCTCCCAACAGCAAGAGAGATACCAAGGCTACATGGGGCTTTAAGGTGTCTGCT 1691

645 GAGGAGCCCATCTACAGGTTATGCTGGAGCCTTGTGAGGTGCTGTGAGACACAATGCTGAGAT 704  
 224 ValAlaPheValLysHisSerThrValLeuGluAsnThrAspGlyLysThrLeuProSer 243  
 705 GTGGCTTTTACCAGGAAGCAGGTATTTGAGGAG-----TTACCAAT 749  
 244 TrpGlyGlnAlaLeuLeuSerGlnAspPheGluLeuLeuCysArgAspGlySerArgAla 263  
 750 AAAGCCGAA-----AGGACACAGTACAAGCTCTCTGCCACACAACCTCGAAG 800  
 264 AspValThrGluTrpArgGlnCysHisLeuAlaArgValProAlaHisAlaValVal 283  
 801 CCGGTGACAGAATACAAAGAGTGCACCTGGGCCAAGTCCCTTCATGCTGTGATATCC 860  
 284 ArgAlaAspThrAsp---GlyGlyLeuIlePheArgLeuLeuAsnGluGlyGlnArgLeu 302  
 861 CGAAGCACGAATGACAAAGAGGCCATCTGGGAGCTTCTCGCCAGTCACAGGAGAAG 920  
 303 Phe---SerHisGluGlySerSerPheGlnMetPheSerSerGluAlaTyrGlyGlnLys 321  
 921 TTTGGAAAAAACAAGCATCGGATTCAGCTCTTTGCTCCCTCCCTCG---GGACAGAAG 977  
 322 AspLeuLeuPheLysAspSerThrSerGluLeuValProIleAlaThrGlnThr---Tyr 340  
 978 GACCTGCTGTTCAAGAGTCTGCCATTTGGCTTTGTGAGGGTTCCCCAGAGGTAGATGTA 1037  
 341 GluAlaTrpLeuGlyHisGluTyrLeuHisAlaMetLysGlyLeu----- 355  
 1038 GGGCTCTACCTGACCTTCAGCTTACCCATCCATACAGAACTGTAATAAAAGACGAG 1097  
 356 -----LeuCysAspProAsnArgLeuProTyrLeuArgTrpCysValLeuSerThr 373  
 1098 GATGTGATAGCTTCAAAAGCCCGG-----GTCACATGTTGCTGCGTGGGCACT 1145  
 374 ProGluIleGlnLysCysGlyAspMetAlaValAlaPheArgArgGlnArgLeuLysPro 393  
 1146 GAGGAGAGCGCAAGTGTGATCAGTGGAAACAGACAGCAAGCAGAGGAGG----- 1193  
 394 GluIleGlnCysValSerAlaLysSerProGlnHisCysMetGluArgIleGlnAlaGlu 413  
 1194 ---GTCACCTGCTATCTATCCACCAGGAGACTGATGTCGTCATCATGAAGGA 1250  
 414 GlnValAspAlaValThrLeuSerGlyGluAspIleTyrThrAlaGlyLysIleTyrGly 433  
 1251 GATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 1307  
 434 LeuValProAlaAlaGlyGluHisTyrAlaProGluAspSerSer----- 448  
 1308 TTAGTTCCAGTCTTGGCAGAAACAGAAATCCTCCAAAGCAATGGCTTGGATTGTGTG 1367  
 449 -----AsnSerTyrTyrValValAlaValAlaValArgArgSerSerHisAla 464  
 1368 AACACACAGCAGGAGGTACCTTCTGTAGCAGCAGTGTAGAGAGAGATGCT---GGC 1424  
 465 PheThrLeuAspGluLeuArgGlyLysArgSerCysHisAlaGlyPheGlySerProAla 484  
 1425 TTCACCTGGAGCTCTTTGAGAGCAAGAAGTCTGCCACACTGCCCTGGACAGGACCGCA 1484  
 485 GlyTTrpAspValProValGlyAlaLeuIleGlnArgGlyPheIleArgProLysAspCys 504  
 1485 GGCTGGAAATCCCTCCATGGGCTGCTGTGTAAC-----CAGACCAAGCTCTGCT 1532  
 505 AspValLeuThrAlaValSerGluPheAsnAlaSerCysValProValAsnAsnPro 524  
 1533 -----AAATTTAATGAGTCTTTAGCCAAAGCTGTGCCCTGGTGTGACCCC 1580  
 525 LysAsnTyrProSerSerLeuCysAlaLeuCysValGlyAspGluGlnGlyArgAsnLys 544  
 1581 AAA-----TCCAATCTCTGTGCCCTGTATTGTTGATGAGAAGGGTGAACAAG 1631  
 545 CysValGlyAsnSerGlnGluArgTyrTyrGlyTyrArgGlyValaPheArgCysLeuVal 564  
 1632 TGTGTCCTCCCAACAGCAAGAGAGATACCAAGGCTACATGGGGCTTTAAGGTGTCTGCT 1691

